Cognitive and Noncognitive Costs of Day Care at Age 0–2 for Children in Advantaged Families

Margherita Fort

University of Bologna, Institute of Labor Economics, and CESifo

Andrea Ichino

European University Institute, University of Bologna, Center for Economic and Policy Research, Institute of Labor Economics, and CESifo

Giulio Zanella

University of Adelaide, University of Bologna, and Institute of Labor Economics

Exploiting admission thresholds to the Bologna day care system, we show using a regression discontinuity (RD) design that one additional day care month at age 0–2 reduces intelligence quotient by 0.5% (4.7% of a standard deviation) at age 8–14 in a relatively affluent population. The magnitude of this negative effect increases with family income. Similar negative impacts are found for personality traits. These findings are consistent with the hypothesis from psychology that children in day care experience fewer one-to-one interactions with adults, with negative effects in families where such interactions are of higher quality. We embed this hypothesis in a model that lends structure to our RD design.

A previous version of this paper circulated under the title "Cognitive and Non-cognitive Costs of Daycare 0–2 for Girls." We are very grateful to the city of Bologna for providing the administrative component of the data set, and in particular we thank Gianluigi Bovini, Franco Dall'Agata, Roberta Fiori, Silvia Giannini, Miriam Pepe, and Marilena Pillati for their invaluable help in obtaining these data and in clarifying the many institutional and administrative details of the admission process and organization of the Bologna Daycare System. We gratefully acknowledge financial support from the Einaudi Institute for Economics and Finance, the European University Institute, the University of Bologna Institute of Advanced Studies,

Electronically published December 12, 2019
[*Journal of Political Economy*, 2020, vol. 128, no. 1]

I. Introduction

Day care for infants and toddlers is a convenient solution for parents who need to return to work soon after the birth of a child. Not surprisingly, enrollment rates in center-based day care are generally growing in countries with a developed labor market.[1] Whether day care 0–2 is also beneficial to children in the long run is less obvious. We study the causal effect of time spent at age 0–2 in the high-quality public day care system offered by the city of Bologna, one of the richest Italian cities,[2] on cognitive and noncognitive outcomes measured at age 8–14. At this age, the short-lived effects of day care 0–2 are likely to have faded away, allowing us to explore longer-term consequences. Identification is based on a regression discontinuity (RD) design that exploits the institutional rules of the application and admission process to the Bologna Daycare System (BDS). This strategy allows us to compare similar children attending day care 0–2, from 4 months up to 36 months, for periods of different length, including no attendance at all, in a context where private day care is almost absent and extended family services are the most relevant substitute for day care.

Fondazione del Monte, Fondazione Rodolfo Debenedetti, Gruppo Hera, and the Ministry of Education, University, and Research (PRIN 2009MAATFS 001). This project would not have been possible without the contribution of Alessia Tessari, who guided us in the choice and interpretation of the psychometric protocols. We also acknowledge the outstanding work of Matteo Escudé, Nurfatima Jandarova, Johanna Reuter, and Zheng Wang (as research assistants); of Valentina Brizzi, Veronica Gandolfi, and Sonia Lipparini (who administered the psychological tests to children); and of Elena Esposito, Chiara Genovese, Elena Lucchese, Marta Ottone, Beatrice Puggioli, and Francesca Volpi (who administered the socioeconomic interviews to parents). Finally, we are grateful to seminar participants at several universities and workshops, as well as to Josh Angrist, Luca Bonatti, Enrico Cantoni, Gergely Csibra, Joe Doyle, Ricardo Estrada, Søren Johansen, David Levine, Salvatore Modica, Cheti Nicoletti, Enrico Rettore, Giovanni Prarolo, and Miikka Rokkanen for very valuable comments and suggestions. Data are provided as supplementary material online.

[1] In the largest member countries of the Organization for Economic Cooperation and Development (OECD) for which data are available, between 2005 and 2016 the average enrollment rate changed from 43.9% to 56.7% in France, from 16.8% (year 2006) to 37.3% in Germany, from 27.3% to 35.5% in Italy, from 16.2% to 22.5% (year 2015) in Japan, from 32.7% to 55.3% in Norway, from 38.2% (year 2010) to 53.4% in South Korea, from 14.9% to 34.8% in Spain, and from 37.0% to 31.5% in the United Kingdom. In the United States, this rate increased from 27.4% in 2006 to 28.0% in 2011. For European Union countries, in 2002 the Barcelona European Council had set a target of 33% of children in day care at age 0–2 (hereafter, “day care 0–2”) by 2010, an objective that was justified as a gender policy. Day care 0–2 is also an expensive form of subsidized early education: in 2013, average public spending per child aged 0–2 in these same countries was (at purchasing power parity) $6,200 in France, $3,400 in Germany, $1,200 in Italy, $3,900 in Japan, $9,600 in Norway, $7,000 in South Korea, $1,400 in Spain, $1,000 in the United Kingdom, and $700 in the United States. Source: OECD Family Database, tabulations PF3.1 (“Public Spending on Childcare and Early Education”) and PF3.2 (“Enrolment in Childcare and Pre-school”).

[2] Bologna, with about 400,000 inhabitants in 2019, is the seventh-largest Italian city and is the regional capital of Emilia Romagna in the north of the country. The day care system that we study is a universal crèche system (*asilo nido*), which in this region is renowned for its high quality even outside the country (Hewett 2001).

Applicants to the BDS provide a preference ordering over the programs for which they are eligible and are assigned to priority groups on the basis of observable family characteristics. Within each priority group, applicants are then ranked on the basis of a household size–adjusted function of family income and wealth (from low to high), which we refer to as the family affluence index (FAI). The vacant capacity of programs in a given year determines FAI thresholds, such that applicants whose FAI is no greater than the threshold of their most preferred program receive an admission offer to that program. Those with a higher FAI either are admitted to a program that they prefer less or in some cases are excluded from all programs. The administrative data that we received from the city of Bologna contain the daily attendance records of each child but no information on outcomes. Thus, between May 2013 and July 2015 we interviewed a sample of children from dual-earner households with cohabiting parents who applied for admission to the BDS between 2001 and 2005 and whose children were between 8 and 14 years of age at the time of the interview. Children were tested by professional psychologists using the fourth edition of the Wechsler Intelligence Scale for Children (WISC-IV) protocol to measure intelligence quotient (IQ) and the Big Five Questionnaire for Children (BFQ-C) protocol to measure the "big five" personality traits. The accompanying parent was interviewed by a research assistant to collect socioeconomic information.

In this affluent population of day care applicants, we find that an additional month in day care 0–2 reduces IQ by about 0.5% on average. At the sample mean (116.4), this effect corresponds to 0.6 IQ points (4.7% of the IQ standard deviation), and its magnitude increases with family income. We also find that for the better-off families in this population, an additional month in day care 0–2 reduces agreeableness and openness by about 1% and increases neuroticism by a similar percentage.

To interpret these findings, we model how children are affected by the decisions of their parents, who face a trade-off between spending time at work, which increases family income and improves child outcomes indirectly, and spending time with their offspring, which enhances child development directly. We allow the trade-off to involve day care programs that may be of a different quality than home care. Moreover, home care may be of a better quality in more affluent households. This hypothesis is supported by a psychological literature emphasizing the importance of one-to-one interactions with adults in child development during the early years of life and that these interactions are more effective if complemented by high human capital and high income.[3] In the BDS setting, the adult-to-child ratio is 1:4 at age 0 and 1:6 at age 1–2 (at the time

[3] See, in particular, Csibra and Gergely (2009, 2011). Other references are reviewed in sec. VII.

our data refer to), while the most frequent care modes when day care 0–2 is not available are parents, grandparents, and nannies, all of which imply an adult-to-child ratio close to one.

The central theoretical insight from the model is that when day care time increases, child skills decrease in a sufficiently affluent household because of the higher quality of home care. However, given the high earning potential of an affluent parent and the possibility to substitute high-quality informal care with the less expensive day care provided by the BDS, the loss of child ability is more than compensated for by an increase of household consumption. Therefore, the affluent parent takes advantage of the offer of the most preferred program even if it decreases child skills, as long as the parent cares enough about household consumption. For a less affluent household, instead, the offer of the most preferred program increases both household consumption and child skills, because home care is of a lower quality than day care. The RD estimand around the FAI thresholds determining whether a child is offered her most preferred BDS program identifies a well-defined weighted average of these heterogeneous effects of day care attendance. Following Card et al. (2015), we show that this estimand can be interpreted as a weighted average of treatment-on-the-treated (TT) effects defined by Florens et al. (2008).

After summarizing the relevant literature in section II, we present the theoretical model in section III and the institutional setting in section IV. Section V describes the interview process to collect child outcomes. Section VI shows how the theoretical model maps into the RD framework and presents our results. Finally, section VII reviews the psychological literature providing support for our interpretation of the evidence, section VIII discusses alternative interpretations, and section IX concludes.

II. Previous Research

This study contributes to the economic literature that investigates how early life experiences shape individual cognitive and noncognitive skills.[4] This literature typically distinguishes between day care 0–2 (e.g., crèches) and childcare at age 3–5 (e.g., preschool/kindergarten programs). Economists devoted considerable attention to the latter, often with a special focus on disadvantaged kids, while paying less attention to the former, especially in more advantaged families.[5]

[4] For recent surveys, see Borghans et al. (2008), Almond and Currie (2011), Heckman and Mosso (2014), and Elango et al. (2016).

[5] Duncan and Magnuson (2013, 127) provide a meta-analysis of the large literature on childcare at age 3–5, concluding that these programs improve children's "pre-academic skills, although the distribution of impact estimates is extremely wide and gains on achievement tests typically fade over time." See also Puma et al. (2012); Carneiro and Ginja (2014);

This is not the case in other disciplines. In a four-decade-old review, Belsky and Steinberg (1978) summarized the findings of day care research in psychology, reporting benefits on standardized measures of intelligence for disadvantaged children but no effects on children from advantaged families and negative effects on noncognitive outcomes across the board. Subsequent reviews by Belsky (1988, 2001) confirmed negative consequences of day care. A central theme in Belsky and Steinberg (1978) is that families are affected in different ways by day care because the latter substitutes for family care of different quality during a developmental stage when adult-child interactions are of paramount importance. Our contribution to this literature is the formalization of this idea in an economic model and its test in a causal framework.

In recent years, economists have devoted more attention to the impact of very early childhood interventions on children's outcomes, reporting mixed results. A first group (Felfe and Lalive [2018] in Germany and Drange and Havnes [2019] in Norway)[6] reports results that apply to a relatively disadvantaged population, finding desirable effects of early day care attendance for both cognitive and noncognitive outcomes, concentrated in particular on girls. On the contrary, Baker, Gruber, and Milligan (2008) found undesirable effects on all types of cognitive and noncognitive outcomes when studying the 1997 universal early day care extension in Quebec (a reform that heavily subsidized day care at age 0–4 in a relatively advantaged population).[7] Kottelenberg and Lehrer (2017) dig deeper into the Quebec data, showing that the negative average estimate hides a positive effect for the less advantaged in that population. Similarly, Duncan and Sojourner (2013, 947) use data from the Infant Health and Development Program in the United States, finding that it "boosted the cognitive ability of low-income children much more than the cognitive ability of higher-income children." In terms of standardized magnitude (for 1 month of attendance), the positive effects found for Germany, Norway, and the United States are about 0.3%, and the negative effects for Quebec are about 0.2%. These sizes are comparable to ours.

Felfe, Nollenberger, and Rodríguez-Planas (2015); Havnes and Mogstad (2015); and Elango et al. (2016). To the best of our knowledge, only Gormley and Gayer (2005), Cascio and Schanzenbach (2013), and Weiland and Yoshikawa (2013) investigate the heterogeneity of effects of childcare at age 3–5 programs by family affluence, finding smaller or zero effects for children in advantaged families.

[6] Precursors of these more recent papers are the Carolina Abecedarian Study (Campbell and Ramey 1994; Anderson 2008), the Milwaukee Project (Garber 1988), and Zigler and Butterfield (1968).

[7] More recently, these authors confirmed the long-run persistence of undesirable effects, with negligible consequences for cognitive test scores and with some of the losses concentrated on boys (Baker, Gruber, and Milligan 2015). Three other recent studies provide indirect evidence consistent with the finding for Quebec by exploiting policy changes that altered the amount of maternal care a child receives at age 0–2: Carneiro, Løken, and Salvanes (2015) for Norway and Bernal and Keane (2011) and Herbst (2013) for the United States.

As for the sign, in line with the Quebec studies, we find a negative effect because our sample and identification provide estimates for relatively affluent families with employed and cohabiting parents in one of the richest and most highly educated Italian cities. This is precisely a context in which the quality of one-to-one interactions at home is likely to be better than the corresponding quality in day care 0–2, even if Bologna is renowned for the high standard of its day care system. Moreover, since girls are more capable than boys of exploiting these interactions in early development (see sec. VII), this is a context in which negative effects for girls should emerge more clearly, and in fact they do in our sample.[8] A second possible reason for the negative sign of our estimate pertains to the characteristics of the day care environment. For instance, both Felfe and Lalive (2018) and Drange and Havnes (2019) study day care settings with an adult-to-child ratio of 1:3. The corresponding ratio at the BDS facilities during the period that we study was 1:4 at age 0 and 1:6 at age 1–2, similar to the prevailing ratio in the Quebec context. In this respect, our study suggests that attention should be paid to the adult-to-child ratio when designing day care 0–2 programs.

Finally, as far as cognitive outcomes are concerned, our negative estimate refers to IQ measured by professional psychologists at age 8–14 (as in Duncan and Sojourner [2013], who use measures of IQ at ages 1–5 and 8), while other studies focus on math and language test scores or on indicators of school readiness (Felfe and Lalive 2018; Drange and Havnes 2019). There is a general consensus that IQ, in addition to being a clinical and standardized indicator, is correlated with a wide set of long-term outcomes, including in particular levels of education, types of occupation, and income (see, e.g., Gottfredson 1997). Currie (2001, 214) notes that the literature on the effects of childcare has shifted toward the use of learning test scores or indicators of school readiness as outcomes, probably because "gains in measured IQ scores associated with early intervention are often short-lived."[9] From this viewpoint, a contribution of our study is to show that day care 0–2 may instead also have long-term negative effects on IQ. As for noncognitive outcomes, our results are in line with Baker, Gruber, and Milligan (2008, 2015) even though, different from them, we focus on the big five personality traits.

[8] Kottelenberg and Lehrer (2014a, 2014b) also focus on the heterogeneity of effects by gender (and age) using data from the Quebec expansion. Effects for girls are also studied, with different results, by Carneiro, Løken, and Salvanes (2015) and Elango et al. (2016). Both positive effects (on emotional regulation, motor skills, and eating) and negative effects (on reasoning and memory) of day care 0–2 in the short run are found by Noboa-Hidalgo and Urzùa (2012) in Chile for children with a disadvantaged background.

[9] The cost of measuring IQ, compared with the increasing availability of almost free administrative data on school outcomes, may contribute to explaining why IQ is used less as an outcome in this literature.

III. Theory

Consider a population of parents who face a trade-off between spending time with their offspring, which enhances child development directly, and spending time at work, which increases household income and so improves child outcomes indirectly.[10] A household is composed of a parent and a child, and there are two periods in life: age 0–2 and post–age 0–2. A parent values household consumption, c, and the ability of the child, θ.[11] The utility function is

$$v(c, \theta) = c + \alpha\theta, \tag{1}$$

where $\alpha > 0$ is the weight of child ability in parental preferences. Two forms of childcare are available: parental childcare and a rationed day care system.[12] Since the parent does not value leisure, she splits the time endowment between work for pay, h, and parental childcare, τ_g, so that a parent's time constraint can be written as $h + \tau_g = 1$. The day care system offers a set of $\mathcal{Z} \geq 2$ programs indexed by $z \in \{0, 1/(\mathcal{Z} - 1), 2/(\mathcal{Z} - 1), \ldots, 1\}$. Vacancies are limited and are allocated via a strategy-proof mechanism that will be described below. Each program z is characterized by a combination of quality, $q_d(z)$, and cost of attendance per unit of time, $\pi_d(z)$. This cost is expressed in units of consumption, and it reflects two components: a transportation cost $k(z)$ and an attendance fee ϕy_{-1}, with $\phi < 1$, that is identical for all programs and is proportional to past household income $y_{-1} = wh_{-1}$, where $w = w(\theta_g)$ is the wage rate (increasing in parental skill θ_g) and $h_{-1} \in [0, 1]$ is past labor supply.[13] Therefore, $\pi_d(z) = k(z) + \phi y_{-1}$. Without loss of generality, we assume that day care programs can be ordered in a way such that the function $s(z) = \alpha q_d(z) - k(z)$ is strictly increasing in z.[14] We later show that, thanks to this assumption, derived utility of parents is also increasing in z, and therefore z is the parents' ranking of programs.

[10] Our framework builds on Becker (1965) as well as Carneiro, Cunha, and Heckman (2003) and Cunha and Heckman (2007). A similar framework is employed by Bernal (2008).

[11] We are indifferent between treating parental preferences over θ as direct—i.e., the parent values child ability per se—or indirect—i.e., the parent values the future earnings of the child, which increase from the child's cognitive or noncognitive skills. We also assume that household consumption benefits both the parent and the child.

[12] The more realistic case that allows for a third type of care acquired from babysitters or within the extended family is considered in the appendix (available online).

[13] In the Bologna context, conditional on a program, parents decide the number of days of attendance but not the number of hours during the day (with few special exceptions). Since every day of attendance requires transportation, total travel cost is proportional to attendance, and this is reflected in our assumption about $k(z)$.

[14] Consistent with the institutional setting described in sec. IV, there are no ties.

Skills are determined at age 0–2 by parental ability, θ_g, household income given by $y = hw$, and the quality of care. Denoting by τ_d time spent by the child in day care, the technology of skill formation is

$$\theta = \eta(\theta_g) + q_g y \tau_g + q_d(z)\tau_d, \tag{2}$$

where $\eta(\cdot) > 0$ captures inherited parental ability, which is also the baseline child ability, and $q_g y$ represents the quality of childcare at home. This specification reflects the idea that while all children attending the same program enjoy the same day care quality $q_d(z)$, the quality of parental care, $q_g y$, differs among children because parental quality q_g is complemented by the cognitive and economic resources of the household, summarized by y. Such complementarity introduces a convexity that ensures that the parent does not necessarily specialize in producing either child quality or income. Equation (2) also indicates that a child would benefit from parental ability θ_g directly, even if the parent had zero earnings.

A child requires a fixed amount of care time $b \in (1/2, 1)$.[15] Therefore, the chosen childcare arrangement must satisfy $\tau_g + \tau_d = b$, so that parental care and day care are perfect substitutes at rate 1 in childcare time but are substitutes at rate $q_d(z)/q_g y$ in child development.

Replacing the time constraints, for each day care program the parent solves

$$\max_{c,\tau_d} c + \alpha\theta \quad \text{s.t.} \quad \begin{cases} c + \pi_d \tau_d = w(1 - b + \tau_d), \\ \theta = \eta(\theta_g) + q_g w(1 - b + \tau_d)(b - \tau_d) + q_d(z)\tau_d, \\ \pi_d = k(z) + \phi y_{-1}, \\ 0 \leq \tau_d \leq b. \end{cases} \tag{3}$$

The key trade-off in this problem is that if the wage rate is greater than the unit cost of day care, then increasing τ_d adds resources for consumption and to complement home care. At the same time, however, it reduces parental time with the child, causing a negative direct effect on child ability if the quality of day care is worse than the quality of parental care.

Let $\underline{z} \geq 0$ be the "reservation program," to be determined below. The parent applies for the subset of programs for which the optimization problem has an interior solution or attains a corner characterized by strictly positive attendance, if the child is offered admission, as well as for the reservation program. Therefore, $A = \{\underline{z}, \dots, 1\}$ is the application set of the parent. Consider first the interior solution for $z \in A$. This is given by

[15] This restriction means that the child needs active care for at least half the time but for less than the entire time (e.g., the parent can work while the child sleeps).

$$\tau_d^*(z) = \frac{2b-1}{2} + \frac{w + \alpha q_d(z) - k(z) - \phi y_{-1}}{2\alpha q_g w}, \quad (4)$$

and parental utility at the optimum is

$$v^*(z) = \tau_d^*[w - \phi y_{-1} + \alpha q_g w(2b - 1 - \tau_d^*) + \alpha q_d(z) - k(z)] + V, \quad (5)$$

where $V = w(1-b)(1 + \alpha q_g b) + \alpha\eta(\theta_g)$. To simplify the analysis, let the number of programs be large enough so that the [0, 1] interval offers a convenient approximation to the set of available programs and z is continuous. Then, using the envelope theorem, $dv^*/dz = (\alpha q_d'(z) - k'(z))\tau_d^*$, and since we have assumed that the ordering of programs implied by z is such that $s(z) = \alpha q_d(z) - k(z)$ is strictly increasing in z, it follows that derived utility $v^*(z)$ must also be strictly increasing in z. Therefore, the condition

$$\alpha q_d'(z) - k'(z) > 0 \quad (6)$$

holds, and the ranking z is consistent with derived preferences over programs.[16]

Admission offers are made on the basis of eligibility thresholds, $\mathcal{Y}_z$. If $y_{-1} \leq \mathcal{Y}_z$, then the child qualifies for program z. For a given application set A, the cutoffs $\mathcal{Y}_z$ faced by a given household are random draws from a distribution that has the same support as the distribution of past household income. Two thresholds are of special interest: $\mathcal{Y}^P \equiv \mathcal{Y}_1$, which determines whether a child is offered her most preferred program ("preferred threshold"), and the maximum threshold in A, $\mathcal{Y}^M \equiv \max_{z\in A}\{\mathcal{Y}_z\}$, which determines whether a child is offered any program ("maximum threshold"). If $y_{-1} > \mathcal{Y}^M$, then the child does not qualify for any of the programs in A and so for all of these programs the problem has a constrained solution $\tau_d^* = 0$. Otherwise, the child is offered the most preferred program among those for which she qualifies.

Our goal is to model how a parent reacts to the offer of the most preferred program $z = 1$ as opposed to the best available alternative. To this end, consider a parent whose y_{-1} is just above $\mathcal{Y}^P$ so that the child barely does not qualify for $z = 1$. The best alternative to the most preferred program may be one of the following:

Case L. A *less* preferred program $z = \ell < 1$, if and only if $\mathcal{Y}^M > y_{-1} > \mathcal{Y}^P$;

Case N. *No offer*, if and only if $y_{-1} > \mathcal{Y}^M = \mathcal{Y}^P$.

[16] As shown in sec. IV.A, eq. (6) is satisfied, on average, in our setting: programs that are ranked higher by parents are typically closer to home, $k'(z) < 0$, and of weakly better quality, $q_d'(z) \geq 0$.

Note that program ℓ is not necessarily in a left neighborhood of 1, and so the ranking difference $1 - \ell$ is a discrete change even if z is continuous. Let $\Delta\tau_d^*$ be the discrete change in optimal day care time induced by the offer of $z = 1$ instead of its best alternative. Given conditions (4) and (6), starting at an interior solution we can establish the following:

REMARK 1 (First stage). The offer of the most preferred program increases day care time. The increase differs in cases L and N:

$$\text{if L then} \quad \Delta\tau_d^* \approx \frac{\alpha q_d'(\ell) - k'(\ell)}{2\alpha q_g w}(1 - \ell) > 0, \tag{7}$$

$$\text{if N then} \quad \Delta\tau_d^* = \tau_d^*(1) - 0 = \frac{2b - 1}{2} + \frac{w + \alpha q_d(1) - k(1) - \phi y_{-1}}{2\alpha q_g w} > 0. \tag{8}$$

Consider now the corner solutions. Remark 1 allows us to characterize the application set of a parent by modeling the possibility that there exists a program $\underline{z} \in [0, 1]$ for which $\tau_d^*(\underline{z}) = 0$. From equation (4), this happens when

$$w - k(\underline{z}) - \phi y_{-1} + \alpha\left[q_g w(2b - 1) + q_d(\underline{z})\right] = 0. \tag{9}$$

If a program satisfying this condition exists, then $\underline{z}$ is the reservation program. If $\underline{z} = 0$, then the application set $A = [\underline{z}, 1]$ coincides with the entire set of programs. If $0 < \underline{z} < 1$, then a corner solution $\tau_d^*(\zeta) = 0$ exists for any $\zeta \leq \underline{z}$. If there is no program for which utility when $\tau_d^* > 0$ is greater than utility when $\tau_d^* = 0$, then the parent does not apply to any program. Finally, there may exist a program $\bar{z} \in (\underline{z}, 1)$, such that $\tau_d^*(\bar{z}) = b$. In this case, for the subset of programs ranked $\zeta \geq \bar{z}$, the optimal day care time is at the $\tau_d^*(\zeta) = b$ corner. Then, if $\bar{z} \leq \ell$, $\Delta\tau_d^* = 0$ in case L and $\Delta\tau_d^* = b > 0$ in case N, and if $\ell < \bar{z} \leq 1$, then $\Delta\tau_d^* = b - \tau_d^*(\ell) > 0$ in case L and $\Delta\tau_d^* = b > 0$ in case N.

Focusing on the cases in which the offer of the most preferred program induces a strictly positive increase in day care attendance, the key prediction of the model concerns the response of child ability at the optimum,

$$\theta^* = \eta(\theta_g) + q_g w(1 - b + \tau_d^*(z))(b - \tau_d^*(z)) + q_d(z)\tau_d^*(z), \tag{10}$$

to an increase in optimal day care time following the offer of $z = 1$. The response differs in cases L and N:

if L then $\Delta\theta^*$

$$\approx \frac{(w - k(z) - \phi y_{-1})k'(\ell) + \alpha^2 q_d'(\ell)\left[q_g w(2b - 1) + q_d(\ell)\right]}{2\alpha^2 q_g w}(1 - \ell), \quad (11)$$

$$\text{if N then } \Delta\theta^* = q_g w(1 - b + \tau_d^*(1))(b - \tau_d^*(1)) + q_d(1)\tau_d^*(1) - q_g w(1 - b)b. \quad (12)$$

Dividing these expressions by equations (7) and (8) gives the treatment effect of interest:

$$\text{if L then } \frac{\Delta\theta^*}{\Delta\tau_d^*} \approx \frac{(w - k(\ell) - \phi y_{-1})k'(\ell) + \alpha^2 q_d'(\ell)\left[q_g w(2b - 1) + q_d(\ell)\right]}{\alpha(\alpha q_d'(\ell) - k'(\ell))}, \quad (13)$$

$$\text{if N then } \frac{\Delta\theta^*}{\Delta\tau_d^*} = \frac{q_g w(1 - b + \tau_d^*(1))(b - \tau_d^*(1)) + q_d(1)\tau_d^*(1) - q_g w(1 - b)b}{\tau_d^*(1)}. \quad (14)$$

These quantities may be positive or negative. However, the following remark holds:

Remark 2 (Treatment effect). Under the conditions specified below, there exist values $\tilde{w}_{\mathrm{L}}$ and $\tilde{w}_{\mathrm{N}}$ of the parental earning potential w, such that if L and if $k'(z) < 0$ and $q_d'(z) \geq 0$ but sufficiently small,[17] then

$$\frac{\Delta\theta^*}{\Delta\tau_d^*} \approx \frac{d\theta^*}{d\tau_d^*} < 0 \Leftrightarrow w > \tilde{w}_{\mathrm{L}} = \frac{k(z)k'(z) - \alpha^2 q_d'(z)q_d(z)}{(1 - \phi h_{-1})k'(z) + (2b - 1)\alpha^2 q_d'(z)q_g} > 0, \quad (15)$$

and if N and if b is sufficiently small,[18] then

$$\frac{\Delta\theta^*}{\Delta\tau_d^*} < 0 \Leftrightarrow w > \tilde{w}_{\mathrm{N}} = \frac{\alpha q_d(1) + k(1)}{1 - \phi h_{-1} + (1 - 2b)\alpha q_g} > 0. \quad (16)$$

In other words, in both cases L and N, the skill response to more day care at the optimum is positive in less affluent households but may be negative in more affluent ones. To see why, note that under the conditions of remark 2, if the parent is sufficiently affluent, then an increase in day care

[17] Specifically, $0 \leq q_d'(z) < -(1 - \phi h_{-1})k'(z)/(2b - 1)\alpha^2 q_g$. Section IV.A suggests that these conditions are satisfied on average in our setting. If $q_d'(z) \geq 0$ and is sufficiently large, then $d\theta^*/d\tau_d^*$ is positive at all levels of affluence.

[18] Specifically, $b < 1/2 + (1 - \phi h_{-1})/2\alpha q_g$. The second term on the right-hand side of this inequality is strictly positive, and it may well be greater than 1/2, in which case no further restriction is imposed on b beyond the assumption that $b \in (1/2, 1)$.

time generates a skill loss because home care is of better quality than day care. However, given the high earning potential of the affluent parent, consumption increases enough with the additional working time to compensate for the skill loss in terms of utility, so that the parent trades off child ability for consumption. For a less affluent parent, in addition to an analogous increase in consumption, there is a gain in terms of child skills because in this case the quality of home care is not sufficiently high compared to the quality of day care.

Whether the skill response to a longer day care exposure is indeed negative for more affluent children is therefore an empirical question. To answer it, we leverage the BDS institutional setting and the associated data, which we describe next.

IV. Institutional Setting

The BDS granted us access to the application, admission, and attendance records for all of the 68 day care facilities operating in Bologna between 2001 and 2005 (of which nine are charter). Every year, these facilities enroll approximately 3,000 children of ages 0, 1, and 2 in full- or part-time modules. Henceforth, we refer to these ages as "grades" and we use the term "program" to define a module (full- or part-time) in a grade (age 0, 1, or 2) of a facility (68 institutions) in a given calendar year (2001–5). There are 941 such programs in our data, and we have information on the universe of 9,667 children whose parents applied for admission to one or more programs of the BDS between 2001 and 2005. Parents can apply to as many programs as they wish in the grade-year combination for which their children are eligible, and they are asked by the BDS to provide a preference ordering of these programs (sec. IV.A). Given these preferences, day care vacancies are allocated by an algorithm that is equivalent to a deferred acceptance (DA) market design (sec. IV.B).[19]

A. *Determinants of Parents' Ranking of Programs*

Table 1 documents that parental preferences over programs systematically reflect distance from home and (to some extent) program quality as measured by their reputation. In the first group of results, geo-referenced information is used to describe the distance in kilometers between each program and the home of the eligible children in the grade-year combination of that program. Mean distance is just above 4 km (SD $\approx$ 2.2), which is also the median distance, and ranges between 100 m and slightly more

[19] See Gale and Shapley (1962) and Roth (2008).

TABLE 1
Distance from Home and Quality of Programs Measured by Their Reputation

	2001	2002	2003	2004	2005
Distance statistics:					
Mean	4.06 [2.21]	4.13 [2.27]	4.09 [2.24]	4.14 [2.27]	4.14 [2.28]
Minimum/ maximum	.02/13.71	.01/14.25	.02/14.10	.02/13.68	.01/14.12
Mean distance from home to:					
Most preferred	1.22 (.04)	1.24 (.04)	1.20 (.03)	1.22 (.03)	1.20 (.03)
Second-most preferred	1.46 (.04)	1.46 (.04)	1.38 (.04)	1.41 (.04)	1.42 (.04)
Third-most preferred	1.66 (.05)	1.71 (.05)	1.57 (.04)	1.68 (.04)	1.66 (.04)
Not most preferred but ranked	1.90 (.02)	1.99 (.02)	1.95 (.02)	2.01 (.02)	1.95 (.02)
Not ranked	4.29 (.01)	4.38 (.01)	4.31 (.01)	4.35 (.01)	4.36 (.01)
Quality statistics:					
Mean	.01 [.26]	.04 [.36]	.03 [.33]	−.01 [.43]	.00 [.42]
Minimum/ maximum	−.86/.65	−1.01/1.00	−1.24/1.40	−1.92/1.02	−1.34/1.23
Mean quality of:					
Most preferred	.07 (.01)	.12 (.01)	.10 (.01)	.10 (.01)	.11 (.01)
Second-most preferred	.06 (.01)	.11 (.01)	.11 (.01)	.09 (.01)	.10 (.01)
Third-most preferred	.06 (.01)	.11 (.01)	.10 (.01)	.03 (.01)	.05 (.01)
Not most preferred but ranked	.05 (.00)	.08 (.01)	.08 (.00)	.03 (.01)	.04 (.00)
Not ranked	.00 (.00)	.03 (.00)	.02 (.00)	−.01 (.00)	−.00 (.00)

Note.—For each year, the table reports statistics on program distance (in km) from the home of applicants and on program quality. The distance measure is based on geo-referenced information. Quality is a reputational indicator constructed in the following way. First, we compute the difference between the average ranking of a program and the average ranking of its alternatives in each grade-year combination for all the households located in a given distance cell from the program and its alternatives. The overall quality of a program is then the average of the distance-specific qualities. Each distance cell is an interval of 0.5 km (an annulus, effectively). Results are based on 5,602 children with two working parents and living within the city boundaries of Bologna. Standard errors are shown in parentheses. Standard deviations are shown in brackets.

than 14 km.[20] The next group of results shows that, on average, the ranking of programs is inversely related to their distance from the home of applicants. The most preferred program is typically located at a distance of 1.2 km. The second- and third-most preferred are located farther away by approximately 200 and 400 additional meters, respectively. The average

[20] These results are based on 5,602 children who have two working parents (i.e., the group on which we focus our study, as explained in sec. IV.B) and live within the city boundaries. For this analysis, we do not consider households living outside the city boundaries because their preferences over programs are probably affected by commuting patterns on which we unfortunately have no information.

distance of programs that are explicitly ranked by parents but that are not their most preferred is slightly less than 2 km, while the most distant programs are those that parents do not rank even if available in their grade-year combination. On average over all programs, moving one position down in the preference ordering is associated with an increased distance of ≈0.35–0.53 km from home. All these differences are statistically significant.

As for quality, we do not have an objective measure, and we rely on a reputational indicator that we constructed in the following way.[21] Consider a set of programs, denoted by j, for which some households, denoted by i and located in a cell of distance d from all these programs, are eligible for. Each distance cell (an annulus, effectively, i.e., a region bounded by two concentric circles) is an interval of 0.5 km up to a maximum of 4 km (the distance beyond which parents typically do not rank programs), so that $d \in \{1, \dots, 8\}$ denotes the eight resulting, nonoverlapping cells of 0.5-km size. Let r_{ijd} be the rank of program j in the application set of household i in distance cell d.[22] The reputation of program j among households in distance cell d is defined as

$$q_{jd} = \bar{r}_{jd} - \bar{r}_{-jd}, \tag{17}$$

where $\bar{r}_{jd}$ is the average ranking of program j in distance cell d, while $\bar{r}_{-jd}$ is the average ranking of the programs different from j in the same cell. Therefore, q_{jd} measures the difference between the average ranking of program j and the average ranking of its alternatives in each grade-year combination for all the households located in the same distance cell d from j and its alternatives. Considering different distance cells, note that each program j is compared with partially different alternatives and by different households in each of these cells, so it may be preferred in some cells but not in others. However, larger values of q_{jd} in different cells imply that j in general has a positive reputation among different groups of households and with respect to different alternatives for given distance.[23] To capture the overall reputation of program j, we compute the average

[21] We do not have information on program-specific teacher-to-children ratios. However, guidelines for programs in the BDS are set at the central level (Comune di Bologna 2010), with little autonomy left to the different facilities. Specifically, the BDS strictly enforces standards concerning goals and daily planning of educational activities and the number of teachers and square meters per child. While programs may still differ, these guidelines suggest a relatively uniform quality across programs. This uniformity is in line with the evidence based on the reputational indicator described below.

[22] For all programs that were not explicitly ranked by a parent, we impute the ranking position that follows the rank of the least preferred among the explicitly ranked programs. This imputation captures the idea that programs not ranked are all indifferently less preferred than the ranked ones. The average fraction of programs not ranked by a parent is about 90% and is constant across years.

[23] The average number of households i, for each combination of program j and distance d, is 138 (108 [SD]) and ranges between 11 (10 [SD]) in cell 1 (from 0 to 0.5 km) and 178 (89 [SD]) in cell 8 (from 3.5 to 4 km).

$$q_j = \frac{1}{8}\sum_{d=1}^{8}\bar{r}_{jd} - \bar{r}_{-jd}. \tag{18}$$

Positive values of q_j indicate a better reputation, meaning that j is systematically more likely to beat its alternatives at all distances. Given the way it is constructed, this measure of quality is centered around zero (third group of results in table 1; SD $\approx 0.3 - 0.4$), but it differs across programs. For example, in 2003, the best program according to this reputational indicator is ranked 1.4 positions better than its alternatives, while the worst program, in 2004, is 1.92 positions worse than its alternatives, on average.

Now consider, as an example, a hypothetical grade-year combination with only three available programs, $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$. If all eligible households unanimously ranked these programs in the same way ($\mathcal{A} \succ \mathcal{B} \succ \mathcal{C}$) at all distances, then their reputation would be ordered as $q_{\mathcal{A}} > q_{\mathcal{B}} > q_{\mathcal{C}}$. In the absence of agreement among households, the reputation of the three programs would instead be similar: $q_{\mathcal{A}} \approx q_{\mathcal{B}} \approx q_{\mathcal{C}}$. The evidence in the last group of results in table 1 suggests that there is little agreement, at least at the top of the rankings. In 2001, 2002, and 2003, there is no statistically significant difference between the average values of q_j for the programs that are ranked in the top positions. Only in 2004 and 2005 is the reputation of the most preferred program (0.10 and 0.11, respectively) significantly larger than the quality of the average less preferred but ranked program (0.03 and 0.04, respectively).

On the basis of this evidence, we conclude that in every year parents certainly prefer programs that are closer to home. As for quality, the revealed reputation of ranked programs shows some convergence of preferences on specific programs in later years, but differences among programs, if they exist, are unlikely to play a major role when parents rank them. Had these differences been of first-order importance, they would have showed up in table 1. Moreover, in the population that we study, 61.6% of applicants are offered their most preferred program and 89.7% receive an admission offer; of these, 84.1% are offered one of their first three choices. In light of these facts, in the remainder of this paper we assume that $q_d'(z) \approx 0$ among the top-ranked programs and in particular at the best alternative $z = \ell$ to the most preferred program $z = 1$.

B. Admission Process

Demand for admission systematically exceeds supply, and there are on average about 1,500 vacancies for about 1,900 applicants each year. The rationing mechanism is based on a lexicographic ordering of applicants. At a first level, applicants to each program are assigned to priority groups on the basis of observable family characteristics. The first group (highest priority) is children with disabilities. Second is children in families assisted by social workers. Third is children in single-parent households, including those resulting from divorce or separation. Fourth is children with two cohabiting and employed parents. Fifth is children in households with two cohabiting parents of whom only one is employed. For brevity, we refer to these priority groups as "baskets" 1–5. At a second level, within each of these five baskets, children are ranked according to an FAI. This is an index of family income and net wealth, adjusted for family size.[24] Families with a lower value of the index (i.e., less affluent families) have higher priority within a basket. The DA algorithm determines for each program a "final" FAI admission threshold, which is defined as the FAI of the most affluent child who receives an offer for that program and accepts it. Given the application set A, these thresholds are effectively random numbers for an applicant. A child qualifies for all programs in A whose final thresholds are greater than the child's FAI and is offered the most preferred program among these.

At the end of the admission process, children can be classified in three mutually exclusive and exhaustive ways: the "admitted and attendants," who have received an admission offer and accepted it; the "reserves," who have not received any offer; and the "waivers," who have received an admission offer and have turned it down. Children who are reserves or waivers in a given year may reapply, be offered admission, and attend day care in later years, as long as they are not older than 2. We consider only the first application of each child. Thus, the possibility to turn down an offer (or to be rejected) and to reapply and attend later is one of the reasons of fuzziness in our RD design. Attending children are charged by the BDS a monthly fee that depends on their FAI but that is independent of actual days of attendance during the month. The fee schedule is well known to potentially interested families before they decide whether to apply,[25] and it is continuous by construction at the admission thresholds that we will use in our design.

C. How FAI Thresholds Can Be Used for the RD Design

To ensure a greater homogeneity of the interview sample (to be described in sec. V), we restrict the entire analysis to children in basket 4 (i.e., children with both parents employed and cohabiting at the time of the application), which is the largest group of applicants (about 70% of

[24] The appendix to sec. IV provides details about how this index is constructed.

[25] The fee is an increasing step function of the FAI. This function increases stepwise along brackets that are about €500 wide, with an initial step (from an FAI of zero) of €17 per month and then constant steps of about €6, before reaching the maximum fee of €400 per month independent of household income. The kink at which the day care fee becomes regressive is located at an FAI of about €30,000, roughly corresponding to a gross annual family income of about €80,000 (all values are expressed in 2010 euros).

the total): 6,770 first applications to 911 programs originate from this basket in the period 2001–5. Of these programs, 80 end up with no vacancies for basket 4 children (i.e., the final FAI threshold is in basket 1, 2, or 3); 285 have sufficient capacity for all basket 4 applicants (i.e., the final FAI threshold is in basket 5), and 546 offer admission to some but not all of the basket 4 applicants (i.e., the final FAI threshold is in basket 4). The remaining 30 (to reach the total of 941 programs) do not receive applications from basket 4 households. Some tables and figures below are based on subgroups of this sample, for the reasons explained in the respective notes and legends. The population of applicants is relatively affluent, particularly in basket 4. The average FAI across the five baskets is about €20,000 (constant 2010 prices), corresponding to a gross annual household income of about €54,000. In basket 4, the average FAI is about €25,000, corresponding to an income of about €67,000. This is roughly twice the average annual gross household income in Italy at the time the data refer to.[26]

In this institutional setting, parents cannot predict final FAI thresholds and thus cannot manipulate their FAI to secure an admission offer. If FAI thresholds were persistent across years, it would be easy for them to find out the final thresholds of the programs they wish to apply for. This is not the case: in a regression of the threshold in year t on the threshold in year $t-1$ (or the most recent previous threshold) for each program, the slope is estimated with high precision to be low (0.14 [0.04 SE]) and R^2 is only 0.02. The top left panel of figure 1 illustrates this lack of persistence. For an accurate guess of FAI thresholds, families would need a formidable amount of additional information.[27] The top right panel similarly shows that for each program, given the number of basket 4 vacancies in year $t-1$, there is substantial variability in the number of vacancies in year t. Moreover, the bottom left panel shows that, given the low degree of socioeconomic segregation across neighborhoods in Bologna, the distribution of thresholds within neighborhoods is almost invariant to changes in average neighborhood income and that the dispersion of these thresholds is large within each neighborhood. Thus, even if parents tried to manipulate their FAI, they would not know by how much the index should be reduced to receive an offer from a specific program.

Additional support for this claim is provided by the continuity of the FAI density and the distribution of pretreatment covariates. In the bottom right panel of figure 1, the density of observations is plotted stacking thresholds and centering them at zero so that the FAI distance from each

[26] Additional descriptive statistics are provided in the appendix to sec. IV.

[27] For example, the vacant capacity of the programs they wish to apply for, the number of applicants to these programs, the FAI of each applicant, how other applicants rank programs, and how many admitted children in each program turn down the offer they receive.

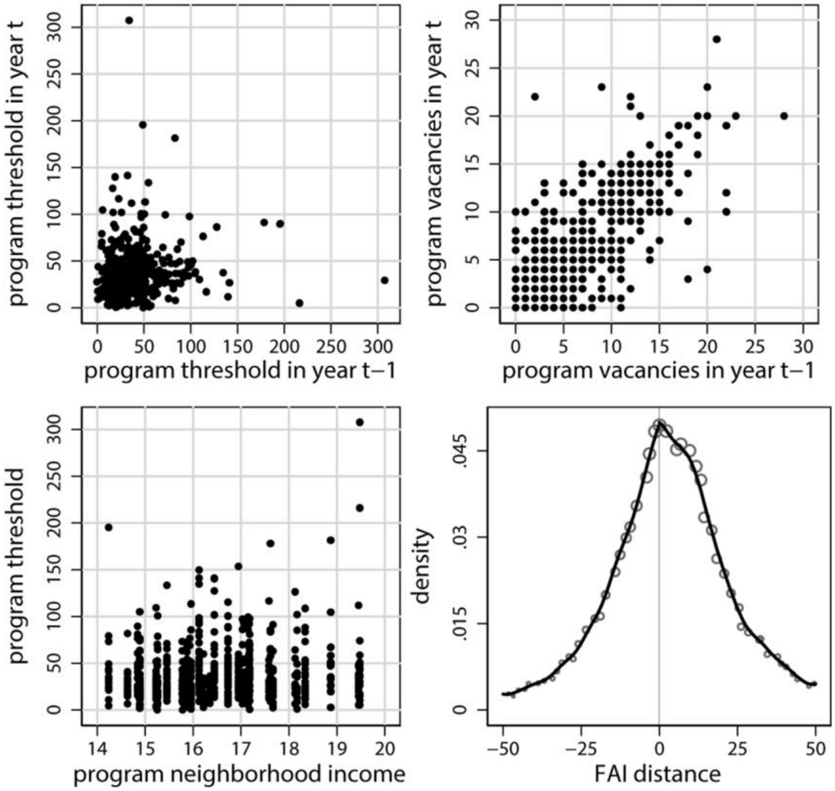

FIG. 1.—Final FAI thresholds, basket 4 vacancies over time, and final FAI density. *Top*, each filled circle represents a program, and the coordinates are either the final FAI thresholds of that program in two consecutive periods (*left*) or the vacant capacity for basket 4 children in two consecutive periods (*right*). Sample: 238 programs with rationing for basket 4 children in two consecutive periods. *Bottom left*, each filled circle represents a program, and the coordinates are the program threshold on the vertical axis and the average income of the program neighborhood on the horizontal axis. *Bottom right*, the open circles represent the frequency distribution inside €2,000 bins (circle size is proportional to population size in the bin), plotted as a function of the distance (thousands of real euros) of a child's FAI from her final FAI threshold ("FAI distance"). The solid lines are from separate local linear regressions on the underlying individual observations on each side of the cutoff, with a triangular kernel and a bandwidth of €5,000. Sample: 5,861 children with two working parents, born between 1999 and 2005, whose parents first applied for admission between 2001 and 2005 to programs with rationing and whose FAI distance from the final FAI thresholds is at most €50,000 and is different from zero.

threshold is the running variable.[28] The McCrary (2008) test rejects the existence of a discontinuity: the gap in the (log) density at the cutoff is

[28] The higher probability mass around the stacked thresholds is because all programs have children immediately to the right and to the left of the cutoff, while children farther away from the cutoff are observed only for programs with a larger number of applications.

-0.007, with a standard error of 0.055. As for pretreatment covariates, the continuity of the distribution of five relevant ones that we observe in the universe (birthday, FAI, average income in the city neighborhood where the program is located, number of siblings at the first application, and number of programs in the application set) is assessed using the test of Canay and Kamat (2018). Results are reported in column 1 of panel A of table 2, and the test never rejects the null that the distribution of any covariate is continuous at the final cutoff in the basket 4 universe.[29]

Given the absence of any evidence of manipulation of the admission process, it would seem natural to use observations around each final FAI threshold for the RD design, but this would be problematic because children applying to many programs would be overrepresented. Specifically, reserve children would appear as many times as the number of programs they apply for, while admitted children and waivers would appear as many times as the number of programs they qualify for. Mapping the model of section III into this institutional setting allows us to circumvent this problem by associating every child with one final FAI threshold only—that is, the unique threshold of her most preferred program, $\mathcal{Y}^P$.[30]

As an example, consider a group of households that apply for the first time in a given year to the same set of five programs whose identity is denoted by $j \in \{\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}\}$. All these parents rank program $\mathcal{C}$ as their most preferred one, so that $z = 1$ for $j = \mathcal{C}$ and $\mathcal{Y}_{j=\mathcal{C}} = \mathcal{Y}_{z=1} = \mathcal{Y}^P$, but they may rank the remaining four programs in different ways. The comparison between the children in this group for whom y_{-1} is barely below $\mathcal{Y}^P$ (and so are offered their preferred program) versus those for whom y_{-1} is barely above $\mathcal{Y}^P$ (and so are not offered their preferred program) provides a quasi-experimental variation that allows us to quantify the predictions of the model. For some households in this group, the best alternative if they do not qualify for program $\mathcal{C}$ is qualification for a less preferred program. This is case L in the model. For the remaining households, $\mathcal{Y}_\mathcal{C}$ is also the maximum threshold in their application set, and so not qualifying for the preferred program implies not qualifying for any program at this first application. This is case N in the model. However, since final FAI thresholds are random draws from the viewpoint of applicants, conditional on the number of applications, there is no self-selection of households in cases L or N.[31]

[29] In the appendix, we provide graphical evidence of the continuity of means of these covariates. A similar graphical analysis is provided in the same appendix for all the remaining instances in which the continuity of covariates is tested in table 2.

[30] Of the 911 programs receiving applications from parents in basket 4, only 890 are listed as the most preferred by at least one family out of 6,575 households with nonmissing FAI information.

[31] Among the 6,575 children in the basket 4 universe with nonmissing FAI information, 4,716 (≈72%) are in case L.

TABLE 2
Continuity of the Distribution of Pretreatment Covariates at the FAI Thresholds

	Final Thresholds, Basket 4 Universe (1)	Preferred Thresholds, Basket 4 Universe (2)	Preferred Thresholds, Interview Sample (3)
A. Pretreatment covariates:			
FAI	.38	.11	.59
Siblings	.62	.96	.23
Preferences	.37	.38	.30
Birthday	.23	.32	.49
Neighborhood income	.72	.41	.92
Father's years of education			.28
Mother's years of education			.53
Father's year of birth			.24
Mother's year of birth			.01
Father self-employed			1.00
Mother self-employed			1.00
Cesarean delivery			.54
Joint test, Cramér–von Mises statistic	.23	.66	.56
Joint test, maximum statistic	.18	.27	.35
B. Invitation, response, and interview rates:			
Invitation of universe	.26	.01	
Response of the invited	.46	.70	
Interview of universe	.08	.64	
Joint test, Cramér–von Mises statistic	.46	.70	
Joint test, maximum statistic	.44	.69	

Note.—The table reports, in col. 1, the p-values from the Canay and Kamat (2018) test of the continuity of the distribution of pretreatment covariates at the final FAI thresholds in the basket 4 universe. In the remaining columns, the p-values are reported for the same test at the preferred FAI thresholds (see sec. IV.C), both in the basket 4 universe (col. 2) and in the interview sample (col. 3; see sec. V). The null hypothesis is that the distribution of the covariate is continuous at the cutoff. Panel A considers pretreatment covariates, and panel B considers the invitation, response, and interview rates to be described in sec. V. Samples: 6,086 children (5,061 children for neighborhood income) in col. 1; 6,300 children (5,247 for neighborhood income) in col. 2; 414 children interviewed out of the invited from the basket 4 universe with no missing values in the covariates included in table 6 (375 for neighborhood income) in col. 3. All children are born between 1999 and 2005, their parents first applied for admission between 2001 and 2005, and their FAI distance from the final FAI thresholds is different from zero. The test is implemented using the rdperm package provided by Canay and Kamat (2018) using the default values chosen by these authors for the number of effective observations used from either side of the cutoff and the number of random permutations.

This example applies to each program that is most preferred by a group of applicants and to the corresponding $\mathcal{Y}^p$ threshold. Figure 2 is constructed by normalizing to zero and pooling these different preferred cutoffs and shows how offer rates, attendance rates, and average days of attendance change in a discontinuous way at these thresholds. The running variable is the FAI distance from the preferred cutoff, with positive

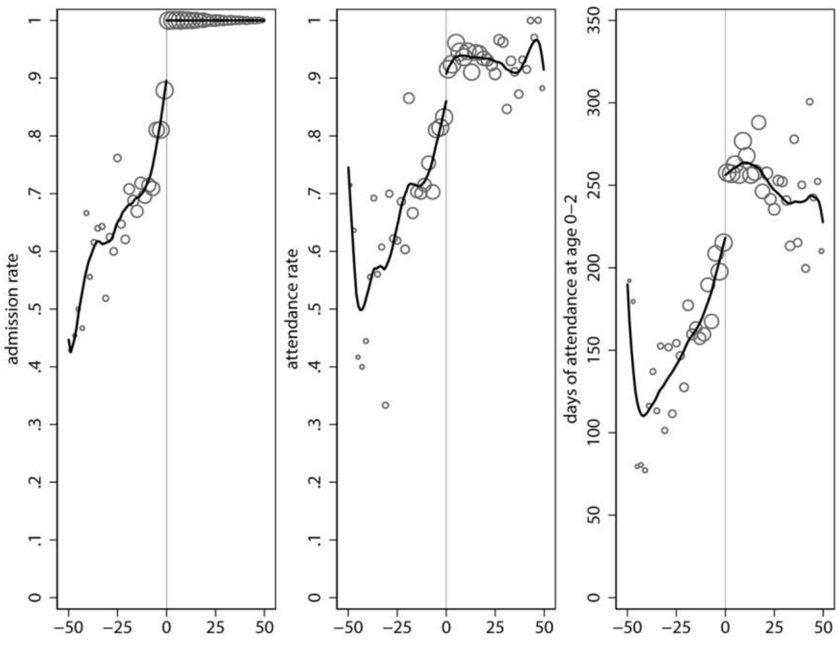

Fig. 2.—Admission offers and attendance around preferred FAI thresholds. The open circles represent offer rates (*left*), attendance rates (*middle*), and average days of attendance at age 0–2 (*right*) inside €2,000 bins, plotted as a function of the distance (thousands of real euros) of a child's FAI from her preferred FAI threshold. The size of a circle is proportional to the number of observations in the corresponding €2,000 bin. The solid lines are from separate local linear regressions on the underlying individual observations on each side of the cutoff, with a triangular kernel and the optimal bandwidth selection developed in Calonico, Cattaneo, and Titiunik (2014); Calonico, Cattaneo, and Farrell (2018); and Calonico et al. (2019) and implemented in STATA by these same authors. Sample: 5,101 children with two working parents, born between 1999 and 2005, whose parents first applied for admission between 2001 and 2005 and whose FAI distance from the final FAI thresholds is at most €50,000 and is different from zero.

values on the right indicating an FAI lower than the threshold. This convention is maintained in all the analogous RD figures that follow. In the left and middle panels, the admission and attendance rates increase sharply (by 10.1 and 4.8 percentage points, respectively) as the FAI crosses the cutoff from higher to lower values, with some fuzziness due to the possibility of reapplying and being offered admission in a later year. These discontinuities translate into a jump of nearly 2 months of attendance (38 working days) in the right panel.[32] On the contrary, the frequency of observations

[32] As expected, given the discussion in sec. III, children in cases L and N attend day care for approximately the same number of months (12.3 and 11.7, respectively) if offered their preferred program. If not offered their preferred program, instead, children in case L attend for about 9 months in a less preferred program, while children in case N attend for about 5.4 months, which is an average between those who reapply and attend in a later year

around the preferred FAI thresholds is continuous. Using the McCrary (2008) test again, the gap in the (log) density at the cutoff is 0.022 with a standard error of 0.13. Similarly, column 2 of panel A of table 2 shows that the Canay and Kamat (2018) test never rejects the null that the distribution of any pretreatment covariate is continuous at the preferred threshold in the basket 4 universe.

Before formalizing this approach to the identification and estimation of the effects of day care 0–2 (sec. VI), we describe in the next section how we collected the cognitive and noncognitive outcomes that we investigate.

V. The Interview Sample

The administrative records do not contain children outcomes at any stage of their development, nor do they contain pretreatment family characteristics beyond the few mentioned above. Therefore, we organized interviews in the field to collect information on outcomes and socioeconomic background for the children included in our final sample.

Between May 2013 and June 2015, we sent invitation letters via certified mail to 1,383 households (of which 1,379 have nonmissing information) with an FAI sufficiently close to final FAI thresholds and that first applied for admission to a program of the BDS during the period 2001–5. At the time of the invitation, children were between 8 and 14 years of age. In these letters, families were given a brief description of the research project and were invited to contact us (either via email or using a toll-free phone number) to schedule an appointment for an interview. Families were informed that participants would receive a gift card worth €50 usable at a large grocery store and bookstore chain. After a few weeks from receipt of the letter, families who had not yet responded were sent a reminder via email or were contacted by telephone. On arrival at the interview site (a dedicated space at the University of Bologna), the child was administered an IQ test (WISC-IV) and a personality test (BFQ-C) by a professional psychologist, and the accompanying parent was interviewed in a separate room by a research assistant to collect socioeconomic information. Overall, each child and the accompanying parent spent about 3 hours at the interview site. Table 3 reports summary statistics of the resulting cognitive and noncognitive test scores. Both scales are normalized by age. The relatively high average IQ of interviewed children (the normalized IQ scale has a mean of 100 for the Italian population of children in the same age range who took the WISC-IV) is in line with the high socioeconomic status of the population under study.

and those who never attend any BDS program. The respective unconditional differences in attendance at the preferred threshold are 3.24 (0.20 [SE]) in case L and 6.30 (0.33 [SE]) in case N.

TABLE 3
IQ and Personality Traits, Summary Statistics

	Mean	Standard Deviation	Median	Minimum	Maximum
IQ	116.4	12.4	116	75	158
Openness	47.7	9.0	48	18	68
Conscientiousness	47.6	10.0	47	16	71
Extraversion	47.4	9.7	48	17	72
Agreeableness	53.4	9.1	54	5	71
Neuroticism	48.3	8.2	47	28	74

Note.—The table reports summary statistics of test scores from the WISC-IV (an IQ test) and the BFQ-C (a personality test) in the interview sample. As explained in the main text, the number of observations are 444 for IQ and 447 for the personality traits (446 for agreeableness). For IQ, only the summary score (full-scale IQ) is considered here; descriptive statistics for the four underlying subscales (verbal ability, working memory, perceptual reasoning, and processing speed) are reported in the appendix to sec. V (available online).

We obtained information for 458 children, corresponding to a response rate of 33.2% of the invited (about 40% in the proximity of final FAI thresholds, as shown in fig. A8; figs. A1–A16 are available online). Of these interviews, only 444 provided a complete set of variables to be used in the econometric analysis when IQ is the outcome and 447 (446 for agreeableness) when the big five personality traits are the outcome.[33] Panel B of table 2 shows, using the Canay and Kamat (2018) test, that our sampling design produces distributions of household invitations, responses, and interviews that are all continuous at the final FAI thresholds. Only for the invitation rate from the universe do we see evidence of a discontinuity at the preferred thresholds. This is not a source of concern in light of the other results reported in table 2, in particular the continuity of the interview rate in the basket 4 universe.

To increase the comparability of children on the two sides of the cutoffs, families were invited starting from those closer to final FAI thresholds. The consequences of this choice are reflected in table 4, which displays the descriptive statistics of key administrative variables for the basket 4 universe, the invited, and the interview samples. As expected, the general pattern is that there are no significant differences between the interviewed and the invited, while both of these groups differ in some dimension with respect to the basket 4 universe. For instance, the invited and the interviewed have a higher FAI than the universe. This is not surprising given how we invited families. The table also shows that the offer rate is higher in the universe than in the interview/invited samples. This happens because, given a large admission rate in the BDS, sampling around

[33] In seven cases, parents informed us that their children had already been tested recently using the WISC-IV, and this test does not provide reliable information if replicated. In seven additional cases, parents did not answer all of the socioeconomic questions, thus generating missing values in some relevant pretreatment variables. For agreeableness, an outlier is not used in the analysis.

TABLE 4
Descriptive Statistics for the Basket 4 Universe, the Invited, and the Interview Samples

Variable	Basket 4 Universe	Invited	Interview	p-Value
FAI at first application	24.87	26.50	27.10	.005
	(20.50)	(19.70)	(17.55)	[.547]
				{.010}
Number of preferences at first application	5.42	5.29	5.59	.199
	(3.66)	(3.42)	(3.53)	[.120]
				{.341}
Siblings at first application	.48	.49	.54	.749
	(.66)	(.65)	(.70)	[.151]
				{.079}
Day of birth in the year	182.8	186.6	180.5	.226
	(104.1)	(106.6)	(111.1)	[.310]
				{.674}
Offered admission at first application	.897	.777	.752	.000
	(.304)	(.417)	(.432)	[.297]
				{.000}
Offered preferred program at first application	.616	.511	.473	.000
	(.486)	(.500)	(.500)	[.169]
				{.000}
Waiver at first application	.124	.075	.068	.000
	(.330)	(.263)	(.251)	[.607]
				{.000}
Year first applied	2003.1	2003.4	2003.5	.000
	(1.43)	(1.42)	(1.38)	[.135]
				{.000}
Year child born	2002.0	2002.5	2002.6	.000
	(1.58)	(1.63)	(1.62)	[.086]
				{.000}
Grade first applied for	.882	.568	.541	.000
	(.786)	(.673)	(.676)	[.459]
				{.000}
Total days of attendance	212.2	223.6	230.5	.010
	(143.3)	(151.4)	(156.3)	[.417]
				{.017}
Ever attended (share with days >0)	.847	.784	.782	.000
	(.360)	(.411)	(.414)	[.916]
				{.001}
Observations	6,575	1,379	444	

Note.—The table compares the means of variables from the administrative records in the basket 4 universe (6,575 children born between 1999 and 2005 whose parents first applied for admission between 2001 and 2005; 6,572 children for "Number of preferences at first application" as a result of missing values), in the sample invited for an interview (1,379 children from this universe), and in the interview sample (444 children interviewed from the universe with nonmissing IQ score and covariates). The p-values in the last column refer to tests of the equality of means for the basket 4 universe vs. the invited (first row), the invited vs. the interviewed (second row, in square brackets), and the basket 4 universe vs. the interviewed (third row, in braces).

final FAI thresholds implies oversampling reserves. As a result, the attendance rate is somewhat unbalanced as well. Similarly, the rate at which parents are offered the preferred program is higher in the universe than in the invitation and interview samples, where it is roughly balanced. Moreover, children in the basket 4 universe are slightly younger, have first applied for higher grades, have spent fewer days in day care, and turn down admission offers at a higher rate than in the invited/interview samples. These are all consequences of the way we selected the invited families, in an attempt to increase the homogeneity of the sample and the comparability around FAI thresholds. However, these differences are not a threat to the internal validity of our RD design, given the continuity of the distribution of covariates and of the density at the thresholds. The number of preferences and the number of children in the household at first application are instead all similar across the three samples.

As for external validity, table 5 compares the means of selected socioeconomic variables available only for the interview sample with the corresponding means for a representative sample of the population of families with two employed parents of young children in large cities of northern Italy. The comparison reveals that the interview sample is by and large

TABLE 5
Interviewed Sample in Comparison to the Northern Italian Population

	Interview Sample	Northern Italy
Child age	11.1	11.1
	(1.6)	(1.9)
Father age	47.5	46.9
	(4.8)	(4.7)
Mother age	45.1	44.9
	(4.1)	(4.7)
Father's years of education	14.2	13.7
	(3.7)	(2.5)
Mother's years of education	15.5	13.9
	(3.2)	(2.4)
Father self-employed	.236	.145
	(.425)	(.355)
Mother self-employed	.106	.087
	(.308)	(.284)
Observations	444	69

Note.—The table compares the means of variables in the interview sample with the corresponding means in the Bank of Italy Survey of Household Income and Wealth (SHIW). From the SHIW, we selected observations to mimic the basket 4 universe of the BDS administrative files in 2001–5. Specifically, we restricted the sample to households with two cohabiting employed parents from the 2000–2006 waves, living in cities of northern Italy with a population of at least 200,000, and who had children age 8–14 between 2013 and 2015. The average child age reported in the table from the SHIW is the average age of the youngest child in these households.

representative of the corresponding Italian population in terms of demographics. However, parents in our sample are slightly more educated and more frequently self-employed. The higher educational attainment of these parents is relevant for the interpretation of our results because it is one of the reasons why, different from other studies, our estimated effects of day care 0–2 refer to children who can enjoy a relatively richer cultural environment at home by Italian standards.

VI. An RD Design for the Effect of Day Care 0–2

A. *Estimand*

Our goal is to identify and estimate the average effect of additional day care attendance on the log of child ability for children attending for τ_d days, which in our context is the average effect of TT defined by Florens et al. (2008),

$$\mathrm{TT}_{\tau_d|y_{-1}}(\tau_d, y_{-1}) \equiv \int \frac{\partial \ln \theta(\tau_d, y_{-1}, u)}{\partial \tau_d}\, dF_{U|\tau_d, y_{-1}}(u), \tag{19}$$

where $F_{U|\tau_d, y_{-1}}(u)$ is the cumulative distribution function (CDF) of individual heterogeneity U conditional on attendance equal to τ_d and FAI equal to y_{-1}.[34] We follow appendix A.2 and remark 3 of Card et al. (2015)[35] to show that a fuzzy RD design around a specific preferred FAI threshold $\mathcal{Y}^P$ can be used to identify a weighted average of the causal effect of interest on the set of children whose most preferred program has this admission threshold, who react to the offer of their most preferred program versus the best alternative, and who all apply for a given number of programs, as exemplified in section IV.C.

Our setting is characterized by unobserved determinants of attendance and by the possibility that a child is in cases L or N, as discussed in section III. To accommodate these features, we write $\tau_d = \tau_d(y_{-1}, \omega, e)$, where e is the realization of the determinants of noncompliance E (possibly correlated with U) and ω is the realization of $\Omega = \mathbb{I}(\mathcal{Y}^P \neq \mathcal{Y}^M)$, where $\mathbb{I}(\cdot)$ is the indicator function. Thus, the RD estimand at this cutoff can be written as

[34] Under our assumptions, around the preferred FAI threshold and conditioning on observable covariates (most notably, the number of applications), child ability at the optimum can be written in compact form as

$$\theta = \eta(\theta_g) + q_g \frac{y_{-1}}{h_{-1}}(1 - b + \tau_d(z))(b - \tau_d(z)) + q_d \tau_d(z) = \theta(\tau_d, y_{-1}, u),$$

where here and in what follows we omit the asterisk that in sec. III denotes values at the optimum.

[35] For the supplemental material of Card et al. (2015), see https://www.econometricsociety.org/sites/default/files/ECTA11224SUPP.pdf.

$$\beta(\mathcal{Y}^P) = \frac{\lim_{y_{-1}\to\mathcal{Y}^{P,r}} \mathbb{E}[\ln\theta(\tau_d(y_{-1},\omega,e),y_{-1},u)|y_{-1}] - \lim_{y_{-1}\to\mathcal{Y}^{P,l}} \mathbb{E}[\ln\theta(\tau_d(y_{-1},\omega,e),y_{-1},u)|y_{-1}]}{\lim_{y_{-1}\to\mathcal{Y}^{P,r}} \mathbb{E}[\tau_d(y_{-1},\omega,e)|y_{-1}] - \lim_{y_{-1}\to\mathcal{Y}^{P,l}} \mathbb{E}[\tau_d(y_{-1},\omega,e)|y_{-1}]}, \tag{20}$$

where $y_{-1}\to\mathcal{Y}^{P,r}$ indicates that the FAI approaches the preferred cutoff from the right and analogously for $y_{-1}\to\mathcal{Y}^{P,l}$ from the left.[36] The appendix to section VI shows that this estimand is a weighted average (over Ω, E, and U) of the causal effect of interest; that is,

$$\beta(\mathcal{Y}^P) = \int \frac{\partial \ln\theta(\tilde{\tau}_d(\mathcal{Y}^P,\omega,e),\mathcal{Y}^P,u)}{\partial\tau_d}\psi(\omega,e,u,\mathcal{Y}^P)\,dF_{\Omega,E,U}(\omega,e,u), \tag{21}$$

where $\tilde{\tau}_d(\mathcal{Y}^P,\omega,e)$ is a value of day care attendance between $\tau_d^r(\mathcal{Y}^P,\omega,e) \equiv \lim_{y_{-1}\to\mathcal{Y}^{P,r}}\tau_d(y_{-1},\omega,e)$ and $\tau_d^l(\mathcal{Y}^P,\omega,e) \equiv \lim_{y_{-1}\to\mathcal{Y}^{P,l}}\tau_d(y_{-1},\omega,e)$—that is, the levels immediately to the right and immediately to the left of $\mathcal{Y}^P$ for given realizations of Ω and E—and where

$$\psi(\omega,e,u,\mathcal{Y}^P) = \frac{(\tau_d^r(\mathcal{Y}^P,\omega,e) - \tau_d^l(\mathcal{Y}^P,\omega,e))(f_{Y_{-1}|\omega,e,u}(\mathcal{Y}^P)/f_{Y_{-1}}(\mathcal{Y}^P))}{\int(\tau_d^r(\mathcal{Y}^P,\omega,e) - \tau_d^l(\mathcal{Y}^P,\omega,e))(f_{Y_{-1}|\omega,e}(\mathcal{Y}^P)/f_{Y_{-1}}(\mathcal{Y}^P))\,dF_{\Omega,E}(\omega,e)}. \tag{22}$$

These weights imply that the children contributing to the estimand are the treated whose attendance changes at the cutoff when they are offered their most preferred program.

To make explicit the presence of cases L and N, equation (21) can be written as

$$\begin{aligned}\beta(\mathcal{Y}^P) &= \int\Bigg[\omega\frac{\partial\ln\theta(\tilde{\tau}_d(\mathcal{Y}^P,1,e),\mathcal{Y}^P,u)}{\partial\tau_d}\psi(1,e,u,\mathcal{Y}^P) \\ &\qquad + (1-\omega)\frac{\partial\ln\theta(\tilde{\tau}_d(\mathcal{Y}^P,0,e),\mathcal{Y}^P,u)}{\partial\tau_d}\psi(0,e,u,\mathcal{Y}^P)\Bigg]\,dF_{\Omega,E,U}(\omega,e,u) \\ &= \mu\mathbb{E}\left[\frac{\partial\ln\theta(\tilde{\tau}_d(\mathcal{Y}^P,1,e),\mathcal{Y}^P,u)}{\partial\tau_d}\psi(1,e,u,\mathcal{Y}^P)\right] \\ &\qquad + (1-\mu)\mathbb{E}\left[\frac{\partial\ln\theta(\tilde{\tau}_d(\mathcal{Y}^P,0,e),\mathcal{Y}^P,u)}{\partial\tau_d}\psi(0,e,u,\mathcal{Y}^P)\right],\end{aligned} \tag{23}$$

[36] Superscripts r and l in $\mathcal{Y}^{P,r}$ and $\mathcal{Y}^{P,l}$ are chosen so to be consistent with the convention adopted in the RD figures above, where we assume that y_{-1} is ordered from higher values on the left to lower values on the right, so that admission to the preferred program occurs to the right of the cutoff $\mathcal{Y}^P$. Note that $\tau_d(\cdot)$ in eq. (20) also depends on what is offered to

where $\mu \equiv \mathbb{E}(\Omega|\mathcal{Y}^P)$—that is, the probability that a child is in case L. This follows from the fact that Ω is stochastically independent of (U, E) given $\mathcal{Y}^P$ because, as argued in sections III and IV.C, FAI thresholds are random draws from the viewpoint of applicants, conditioning on the number of applications. Therefore, there is no self-selection in cases L and N so that μ is an exogenous probability at a given $\mathcal{Y}^P$.

In other words, our estimand at a specific preferred FAI threshold can also be interpreted as an average, weighted by the probability that a child is in case L or N, of the weighted averages of the causal effects of interest for children in each of these two cases. These averages are evaluated at the levels of day care attendance $\tilde{\tau}_d(\mathcal{Y}^P, 1, e)$ and $\tilde{\tau}_d(\mathcal{Y}^P, 0, e)$ that are located between the values of attendance immediately to the right and immediately to the left of $\mathcal{Y}^P$ for each realization of e in cases L and N.

Integrating over the different preferred FAI thresholds that characterize our institutional setting, we obtain the overall average of the $\beta(\mathcal{Y}^P)$'s, weighted by the frequency of observations attached to each cutoff:

$$\beta = \int \beta(\mathcal{Y}^P)\, d\mathcal{F}(\mathcal{Y}^P), \tag{24}$$

where $\mathcal{F}$ is the distribution of preferred FAI thresholds.[37] This solution to the aggregation problem is in the spirit of Cattaneo et al. (2016) and is also implemented by Card et al. (2015), where multiple cutoffs arise from pooling different years.[38]

The estimand β in equation (24) is relevant not only for parents but also for a policy maker interested in expanding vacancies in the existing facilities. As a consequence of this expansion, a larger number of households would have access to their preferred program. Our estimates speak

the parent on the two sides of the cutoff—i.e., $z = 1$ on the right $(y_{-1} \rightarrow \mathcal{Y}^{P,r})$ and $z = \ell$ or no offer on the left $(y_{-1} \rightarrow \mathcal{Y}^{P,l})$. To simplify the notation, we do not make this dependence explicit in $\tau_d(\cdot)$, although it is taken into account in the derivations that follow, as indicated by the notation $\tau_d^r(\cdot)$ and $\tau_d^l(\cdot)$ introduced below.

[37] Equation (24) implicitly assumes that the frequency distribution of the cutoffs over the population of households located just to the right of them is identical to the analogous distribution to the left. To support this assumption, in the appendix to sec. VI we consider the empirical distribution functions of the preferred FAI thresholds in the right and left neighborhoods, and we test their similarity using the Wilcoxon rank-sum test. The null hypothesis of similarity cannot be rejected, both in the basket 4 universe (p-value: .21) and in the interview sample (p-value: .41). Figure A14 plots the two pairs of CDFs.

[38] The application in Card et al. (2015) is the effect of unemployment benefits on unemployment duration in Austria. In their setting, the running variable is annual earnings, and the earnings cutoff at which the benefit schedule has a kink varies by year. Moreover, different from us, their running variable measures with error the underlying assignment variable used to determine unemployment benefits. In our application, instead, the FAI used by the BDS to allocate day care is exactly observed. As shown in the appendix to sec. VI, the absence of measurement error in our setting simplifies the econometric model with respect to Card et al. (2015).

precisely about the effect of such a policy, which may have undesirable consequences on the ability of more affluent children.[39]

B. Empirical Model

Let θ_i be the skill trait of child i, observed at age 8–14, and let $\tau_{d,i}$ denote the treatment intensity, measured as months spent in day care over the entire age 0–2 period.[40] The running variable is the FAI, y_i, at first application, and the equation we estimate is[41]

$$\ln \theta_i = \alpha + \beta\tau_{d,i} + m(y_i) + \gamma A_i + \delta X_i + \epsilon_i, \tag{25}$$

where β is an empirical counterpart of the theoretical estimand derived in equation (24), $m(y_i)$ is a second-order polynomial in the running variable, A_i is a vector of variables describing the application set of a child (specifically, the number of programs included in the application set and dummies for the city neighborhood of the preferred program), and X_i is a vector of pretreatment variables (parents' education, parents' year of birth, number of siblings at the first application, whether parents were self-employed—as opposed to employees—during the year preceding the first application, birthday, and a dummy for cesarean delivery of the child).[42] Finally, ϵ_i captures unobserved determinants of ability.

As usual in RD designs, the inclusion of pretreatment variables is not strictly necessary for identification, but it may increase efficiency, and most importantly, similar estimates of β when observables are included or not support the validity of the identifying assumption that pretreatment covariates are continuous at the thresholds (Imbens and Lemieux 2008; Lee and Lemieux 2010). More direct evidence on the validity of this assumption in the interview sample is provided in column 3 of panel A of table 2, which shows that the Canay and Kamat (2018) test does not reject the null that the distributions of pretreatment covariates are jointly continuous at the preferred FAI threshold.[43]

[39] It is true that other weighting schemes—e.g., those discussed by Bertanha (2017)—would be informative on more general counterfactual scenarios in other institutional settings. An analysis in this spirit is presented in the appendix to sec. III, where we simulate a calibrated version of the theoretical model under alternative weighting schemes.

[40] In the administrative data, we observe the precise daily attendance of children in day care. For convenience, we rescale days of attendance in months defined as 20 working days.

[41] In this parametric specification, we do not center and stack thresholds, different from what we do in the continuity test, thus avoiding the problems generated by observations located precisely at the thresholds. In the appendix to sec. VI, we also report nonparametric estimates that confirm the general pattern of the results described below, although those pertaining to noncognitive outcomes are less precise.

[42] Descriptive statistics for these variables are reported in tables A7 and A8; tables A1–A30 are available online.

[43] As a result of the information acquired from parents in the interviews, here we can assess continuity for a set of 11 covariates, which is larger than the one observed in the universe. In only one case—mother's year of birth—the p-value is less than 5%.

We estimate equation (25) by instrumental variables (IV) using as an instrument a dummy P_i indicating whether a child qualifies for her most preferred program at the first application,

$$P_i = \mathbb{I}(y_i \le \mathcal{Y}_i^P), \tag{26}$$

where $\mathcal{Y}_i^P$ denotes the FAI threshold of child i's most preferred program.

Figure 3 replicates for the interview sample the evidence of figure 2, which was based on the basket 4 universe. Also in this sample, the admission rate, attendance rate, and days of attendance all jump discontinuously at the preferred thresholds (by 19.2 percentage points, 3.9 percentage points, and 41.3 days, respectively).

Monotonicity of the treatment in the instrument must also be satisfied for β to identify the estimand in equation (24). Remark 1 shows that we

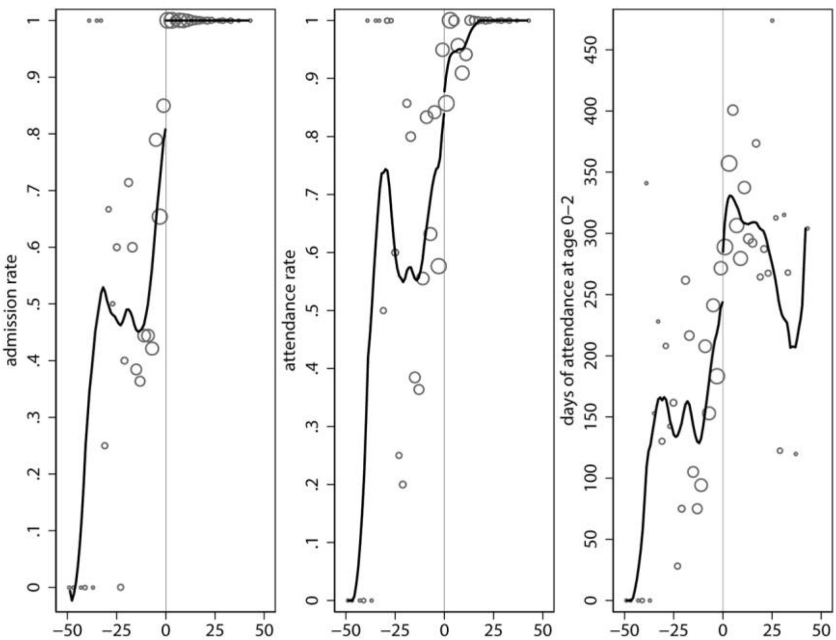

FIG. 3.—Admission offers and attendance around preferred FAI thresholds, interview sample. The open circles represent offer rates (*left*), attendance rates (*middle*), and average days of attendance at age 0–2 (*right*) inside €2,000 bins, plotted as a function of the distance (thousands of real euros) of a child's FAI from her preferred FAI threshold. The size of a circle is proportional to bin size. The solid lines are from separate local linear regressions on the underlying individual observations on each side of the cutoff, with a triangular kernel and the optimal bandwidth selection developed in Calonico, Cattaneo, and Titiunik (2014); Calonico, Cattaneo, and Farrell (2018); and Calonico et al. (2019) and implemented in STATA by these same authors. Sample: 373 interviewed children with two working parents, born between 1999 and 2005, whose parents first applied for admission between 2001 and 2005 and whose FAI distance from the final FAI thresholds is at most €50,000 and is different from zero.

should expect monotonicity to hold in our setting: the offer of the most preferred program unambiguously implies that day care attendance increases weakly for all parents and strictly so for at least some. This prediction is supported by the evidence of figure 4. In the left panel, following Angrist and Imbens (1995), we plot the CDF of days of attendance for the two groups of children defined by the instrument. Visual inspection indicates that the distribution of days of attendance for those who are offered their most preferred program (solid line) first-order stochastically dominates the corresponding one for those who are not (dashed line). Under local "type independence" (Fiorini and Stevens 2017), this is a necessary condition for monotonicity.[44] We use the procedure developed by Barrett and Donald (2003) to formally test this ordering, and we cannot reject the null (p-value: .9998; for full details, see table A10). The right panel of figure 4 plots the estimates of the effect of being offered the most preferred program at different quantiles of months of attendance, on the basis of our preferred specification with all the controls. These estimates are always positive and statistically significant, suggesting no violation of monotonicity also conditional on covariates.

C. Results: Cognitive Skills

The first row of panel A in table 6 reports estimates of the effect of barely qualifying for the most preferred program on IQ, which we refer to as the intention to treat (ITT) effect.[45]

Column 1 includes only the polynomial in the running variable, column 2 adds the application set characteristics, and column 3 includes all controls, yielding similar estimates. Taking column 3 as the preferred specification, the estimated ITT indicates that barely qualifying for the preferred program reduces IQ by 3% (p-value: .005). First-stage estimates are reported in the second row of the table,[46] and they indicate that barely qualifying for the preferred program increases attendance by about

[44] In our context, local type independence requires that the joint distribution of days of attendance in case of qualification for the most preferred program or in case of no qualification (potential treatments) is independent of the running variable (FAI) locally at the cutoff. This assumption is not needed for the identification of the TT in eq. (19) (see the appendix to sec. VI and the discussion in Cattaneo, Frandsen, and Titiunik [2015] and de la Cuesta and Imai [2016]), but we maintain it, as conventional in the literature, to test for monotonicity.

[45] Specifically, we estimate the following equation,

$$\ln \theta_i = \tilde{\alpha} + \tilde{\beta} P_i + \tilde{m}(y_i) + \tilde{\gamma} A_i + \tilde{\delta} X_i + \tilde{\epsilon}_i,$$

where $\tilde{m}(y_i)$ is a second-order polynomial in the FAI and $\tilde{\beta}$ is the ITT effect.

[46] In this case, we estimate the first-stage equation,

$$\tau_{d,i} = \bar{\alpha} + \bar{\beta} P_i + \bar{m}(y_i) + \bar{\gamma} A_i + \bar{\delta} X_i + \bar{\epsilon}_i,$$

where $\bar{m}(y_i)$ is a second-order polynomial in the FAI and $\bar{\beta}$ is the first-stage estimate.

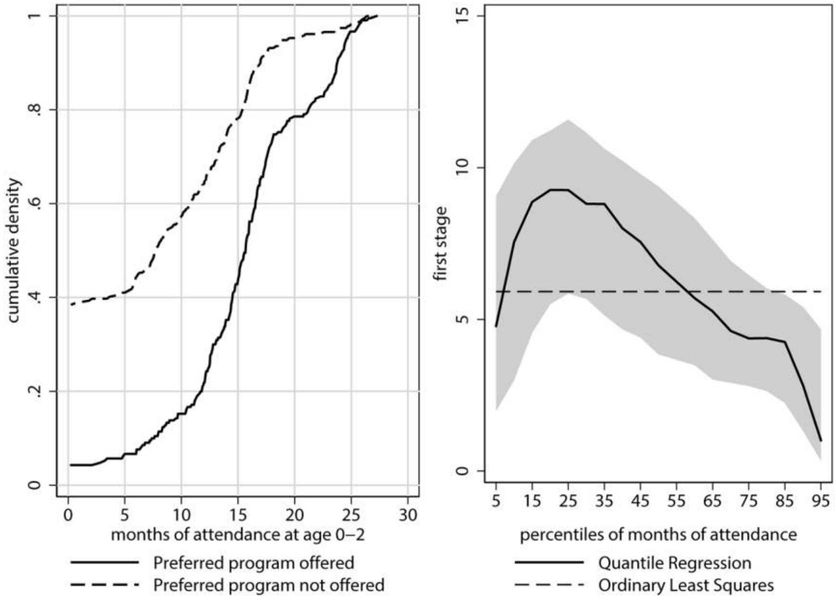

FIG. 4.—Monotonicity of the treatment in the instrument. *Left*, CDF of days of attendance in day care 0–2 for the two groups of children defined by the instrument (whether the child qualifies for the preferred program). *Right*, coefficients from quantile regressions of total days of attendance in day care 0–2 on the instrument and the same controls included in the estimation of equation (25). The running variable is the FAI, and the polynomial in the running variable is of second order. The shaded areas represent the 95% confidence intervals based on 1,000 block-bootstrap replications (to preserve dependence within programs). Each coefficient is obtained by running a separate quantile regression for the 19 quantiles from 0.05 to 0.95. The dashed, horizontal line represents the corresponding first-stage ordinary least squares estimates. Sample: 444 interviewed children with two working parents and nonmissing IQ score and covariates, born between 1999 and 2005, whose parents first applied for admission between 2001 and 2005.

6 months. The *F*-test statistic on the excluded instrument indicates that weak instruments are not a concern.

Rescaling the ITT effect by the first stage gives the IV estimate of the effect of one additional month in day care. In our preferred specification (and similarly in the others), this is a statistically significant loss of about 0.5% (*p*-value: .004), which at the sample mean (116.4) corresponds to 0.6 IQ points and 4.7% of the IQ standard deviation. As argued in section V, the interview sample is characterized by relatively affluent and educated parents in one of the richest Italian cities. Therefore, in light of remark 2, it should not come as a surprise that the IQ effect of day care turns out to be negative in this population.

To further illustrate the empirical relevance of remark 2, we split children into two groups according to whether the preferred FAI threshold they are associated with is above or below the median of all preferred

TABLE 6

EFFECTS OF DAY CARE 0–2 ATTENDANCE ON LOG IQ, FOR ALL CHILDREN AND BY LEVEL OF THE PREFERRED FAI THRESHOLD

	A. ALL FAI THRESHOLDS (Mean Threshold: €24,600)			B. FAI THRESHOLDS ≤ MEDIAN (Mean Threshold: €16,400)			C. FAI THRESHOLDS > MEDIAN (Mean Threshold: €33,000)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ITT effect of qualifying for the preferred program	−.026*	−.030**	−.030**	−.010	−.018	−.018	−.044*	−.045*	−.048**
	(.010)	(.011)	(.010)	(.016)	(.015)	(.014)	(.017)	(.018)	(.017)
First-stage effect of qualifying on months of attendance	6.3**	6.4**	5.9**	5.6**	5.6**	5.1**	5.5**	5.8**	5.6**
	(.9)	(.9)	(.9)	(1.4)	(1.5)	(1.5)	(1.3)	(1.3)	(1.2)
IV effect of 1 month of day care attendance	−.004**	−.005**	−.005**	−.002	−.003	−.003	−.008*	−.008*	−.009**
	(.002)	(.002)	(.002)	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)
F-statistic on excluded instruments	49.1	46.9	44.8	15.2	14.3	11.6	19.6	19.0	19.3
Observations	444	444	444	224	224	224	220	220	220
Polynomial in FAI	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Application set controls		Yes	Yes		Yes	Yes		Yes	Yes
Pretreatment controls			Yes			Yes			Yes

NOTE.—The table reports parametric estimates of the effect of 1 month of day care 0–2 on log IQ, and the associated ITT and first stage, at all levels of the preferred FAI threshold and separately for preferred FAI thresholds below or above the median preferred FAI threshold. ITT coefficients are from regressions of log IQ on the instrument (whether the child qualifies for the preferred program) and controls (see n. 46). First-stage coefficients are from regressions of months (1 month = 20 days of attendance) spent in day care 0–2 on the instrument and controls (see n. 45). IV coefficients are from regressions of log IQ on months of attendance and controls using a dummy for qualification in the preferred program as the instrument (see eq. [25]). The running variable is the FAI, and the polynomial in the running variable is of second order. Sample: 444 interviewed children with two working parents, born between 1999 and 2005, with nonmissing outcome or covariates and whose parents first applied for admission to day care between 2001 and 2005. Robust standard errors are shown in parentheses, clustered at the facility level.

* Significant at 5%.

** Significant at 1%.

thresholds. Results are reported in panels B and C of table 6.[47] For the less affluent group, the estimates refer to the effect of day care 0–2 around a preferred FAI threshold of €16,400 on average (corresponding to a gross annual family income of about €43,000), while in the more affluent group above the median the average threshold is €33,000 (annual family income of about €88,000). Parametric estimates of the IQ response to more day care in these two groups are reported in panels B and C of table 6. In the less affluent group, the ITT effect of barely qualifying for the most preferred program is estimated to be negative but relatively small (less than 2%) and statistically insignificant. In the more affluent group, the estimated loss is nearly three times larger (almost 5%) and precisely estimated. The second row in these same panels displays the first-stage effect, which is similar for more and less affluent households in our sample. Rescaling the ITT effect by the first stage gives a statistically significant IV estimate for the IQ loss of between 0.8% and 0.9% in the more affluent group, while for less affluent households the estimate is an insignificant 0.2%–0.3%. A similar pattern emerges from the nonparametric estimates for the two groups reported in the appendix to section VI, where we also summarize the analysis for the four subscales that compose the total IQ score considered here. With different degrees of intensity, these results hold similarly for the subscales.

As indicated by equation (23), the IV estimates reported in table 6 are averages of the effects of one additional month in day care for children in case N—$\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M) = 1 - \Omega$ (32% in the interview sample)—and for the remaining children, who are in case L: $\mathbb{I}(\mathcal{Y}^P \neq \mathcal{Y}^M) = \Omega$. Table 7 reports results from a specification that allows the IV effects to be different in the two cases. Starting with column 1, which is based on the entire interview sample, the IV estimate of the IQ loss induced by one additional month in day care attendance is 0.7% for children in case L, with no significant difference detected for the remaining children in case N.[48] When

[47] The appendix reproduces the figure and tables of the main text related to the validity of our identification strategy separately for the two affluence groups.

[48] Specifically, we estimate by IV

$$\ln\theta_i = \alpha_{\rm L} + (\alpha_{\rm N} - \alpha_{\rm L})(1 - \Omega_i) + \beta_{\rm L}\tau_{d,i} + (\beta_{\rm N} - \beta_{\rm L})\tau_{d,i}(1 - \Omega_i) + m(y_i) + \gamma A_i + \delta X_i + \epsilon_i,$$

where $\beta_{\rm L}$ is the estimate for case L and $(\beta_{\rm N} - \beta_{\rm L})$ measures how much the estimate for case N differs from that of case L. The instruments are P_i and $P_i(1 - \Omega_i)$. Like in the basket 4 universe (see n. 32) and as expected given the discussion in sec. III, children in cases L and N attend day care for approximately the same number of months (16 and 15, respectively) if offered their preferred program. If not offered their preferred program, instead, children in case L attend for about 9 months in a less preferred program, while children in case N attend for about 5 months, which is an average between those who reapply and attend in a later year and those who never attend any BDS program. The respective first stages for day care attendance at the preferred threshold are 3.93 (0.82 [SE]) in case L and 9.91 (0.62 [SE]) in case N. Additional descriptive statistics for the interview sample children in the two cases are reported in table A21.

TABLE 7
IV Effects of Day Care 0–2 Attendance on IQ, by Cases L and N

	All FAI Thresholds (Mean Threshold: €24,600) (1)	FAI Thresholds ≤ Median (Mean Threshold: €16,400) (2)	FAI Thresholds > Median (Mean Threshold: €33,000) (3)
Day care attendance	−.007**	−.003	−.014**
	(.003)	(.004)	(.005)
Day care attendance × $\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M)$	.004	−.000	.009
	(.003)	(.005)	(.006)
$\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M) \equiv (1 - \Omega)$	−.028	.019	−.098
	(.040)	(.049)	(.077)
Observations	444	224	220

Note.—The table reports parametric estimates of the effect of one additional month of day care 0–2 on log IQ, at all levels of the preferred FAI thresholds and separately for preferred FAI thresholds below or above their median. Coefficients are from regressions of log IQ on months of attendance, months of attendance interacted with $\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M)$ and the full set of controls A_i and X_i, using the dummy P_i for qualification in the preferred program and the same dummy interacted with $\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M) = 1 - \Omega$ as the instruments (see n. 48). The running variable is the FAI, and the polynomial in the running variable is of second order. Sample: 444 interviewed children with two working parents, born between 1999 and 2005, with nonmissing information, whose parents first applied for admission to day care between 2001 and 2005. Robust standard errors are shown in parentheses, clustered at the facility level.

** Significant at 1%.

we split the interview sample by household affluence as in table 6, for the less advantaged children the effect is again small in both cases L and N, while for the more advantaged the IQ loss is confirmed to be large in magnitude (about 1.4% and statistically significant), again with no difference between cases L and N.

D. *Results: Noncognitive Skills*

The corresponding results in the personality domain are reported in tables 8 and 9, where the big five personality traits are the outcome variables and where estimates are reported for all children and by level of the preferred FAI threshold. First-stage estimates are essentially the same as those reported in table 6 and so are not repeated here.[49]

Referring to the preferred specification in panel C, column 9, table 8 shows that qualifying for the preferred day care 0–2 program reduces openness and agreeableness at age 8–14 by 8% and 6.8%, respectively, and increases neuroticism by 5.1% in the more affluent group. No significant

[49] They are not numerically identical, because for these outcomes we have 447 (446 for agreeableness) children instead of 444 (see sec. V and specifically n. 33). All the descriptive statistics are essentially unchanged for this slightly larger sample.

TABLE 8

ITT Effect of Qualifying for the Preferred Program on Personality, for All Children and by Level of the Preferred Threshold

	A. All FAI Thresholds (Mean Threshold: €24,700)			B. FAI Thresholds ≤ Median (Mean Threshold: €16,400)			C. FAI Thresholds > Median (Mean Threshold: €33,000)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Openness	−.032	−.029	−.027	−.002	−.005	.004	−.076*	−.069*	−.080*
	(.021)	(.021)	(.021)	(.040)	(.040)	(.038)	(.029)	(.029)	(.031)
Conscientiousness	.001	.006	.003	.033	.039	.038	−.009	−.007	−.006
	(.023)	(.024)	(.023)	(.035)	(.032)	(.033)	(.037)	(.039)	(.038)
Extraversion	−.032	−.032	−.039+	−.052	−.054	−.055	−.030	−.030	−.037
	(.023)	(.023)	(.022)	(.040)	(.039)	(.037)	(.027)	(.030)	(.032)
Agreeableness	−.034	−.030	−.024	.001	.007	.017	−.060*	−.068*	−.068*
	(.021)	(.022)	(.020)	(.033)	(.033)	(.033)	(.026)	(.033)	(.032)
Neuroticism	.018	.014	.013	−.011	−.013	−.026	.045	.046+	.051+
	(.019)	(.019)	(.020)	(.031)	(.033)	(.033)	(.027)	(.027)	(.027)
Observations	447	447	447	225	225	225	222	222	222
Polynomial in FAI	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Application set controls		Yes	Yes		Yes	Yes		Yes	Yes
Pretreatment controls			Yes			Yes			Yes

Note.—The table reports parametric estimates of the effect of qualifying for the most preferred day care program on the log of scores in the BFQ-C, at all levels of the preferred FAI threshold and separately for preferred FAI thresholds below or above the median preferred FAI threshold. The coefficients are from distinct regressions of each outcome on a dummy for qualification in the preferred program and controls. The running variable is the FAI, and the polynomial in the running variable is of second order. Sample: 447 (446 for agreeableness) interviewed children with two working parents, born between 1999 and 2005, with nonmissing outcome or covariates and whose parents first applied for admission to day care between 2001 and 2005. Robust standard errors are shown in parentheses, clustered at the facility level.

+ Significant at 10%.

* Significant at 5%.

TABLE 9

EFFECTS OF DAY CARE 0–2 ATTENDANCE ON PERSONALITY, FOR ALL CHILDREN AND BY LEVEL OF THE PREFERRED FAI THRESHOLD

	All FAI Thresholds (Mean Threshold: €24,700)			FAI Thresholds ≤ Median (Mean Threshold: €16,400)			FAI Threshold > Median (Mean Threshold: €33,000)		
Openness	−.005	−.005	−.004	−.000	−.001	.001	−.013*	−.011*	−.014**
	(.003)	(.003)	(.003)	(.007)	(.007)	(.007)	(.006)	(.005)	(.005)
Conscientiousness	.000	.001	.000	.006	.007	.007	−.002	−.001	−.001
	(.004)	(.004)	(.004)	(.006)	(.006)	(.006)	(.006)	(.006)	(.006)
Extraversion	−.005	−.005	−.006+	−.009	−.009	−.011	−.005	−.005	−.006
	(.004)	(.003)	(.004)	(.007)	(.007)	(.007)	(.005)	(.005)	(.005)
Agreeableness	−.005	−.005	−.004	.000	.001	.003	−.010*	−.011*	−.012*
	(.003)	(.003)	(.003)	(.006)	(.005)	(.006)	(.005)	(.006)	(.006)
Neuroticism	.003	.002	.002	−.002	−.002	−.005	.008	.007	.009+
	(.003)	(.003)	(.003)	(.005)	(.006)	(.006)	(.005)	(.005)	(.005)
Observations	447	447	447	225	225	225	222	222	222
Polynomial in FAI	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Application set controls		Yes	Yes		Yes	Yes		Yes	Yes
Pretreatment controls			Yes			Yes			Yes

NOTE.—The table reports parametric estimates of the effect of 1 month (20 days of attendance) of day care 0–2 on the log of scores in the BFQ-C, at all levels of the preferred FAI threshold and separately for preferred FAI thresholds below or above the median preferred FAI threshold. The coefficients are from distinct regressions of each outcome on months of attendance and controls using a dummy for qualification in the preferred program as the instrument. The running variable is the FAI, and the polynomial in the running variable is of second order. Sample: 447 (446 for agreeableness) interviewed children with two working parents, born between 1999 and 2005, with nonmissing outcome or covariates and whose parents first applied for admission to day care between 2001 and 2005. Robust standard errors are shown in parentheses, clustered at the facility level.

+ Significant at 10%.

* Significant at 5%.

** Significant at 1%.

effects are detected in this group for either conscientiousness or extraversion, although the point estimates are similarly negative. In the less affluent group, instead, no effects on personality emerge: point estimates are generally closer to zero and often positive (negative for neuroticism). These magnitudes and patterns are in agreement with the analogous ITT effect estimated for IQ in the more affluent group (a significant −4.8%) and in the less affluent one (an insignificant −1.8%).

The finding that for more affluent children sufficient precision is attained for only three of the big five personality traits may be because, in a relatively small sample, the measurement of IQ, being task based, may be more precise than the measurement of personality, which is based on a questionnaire. An alternative explanation is suggested by the sample correlation between IQ and, respectively, openness (0.294), conscientiousness (0.001), extraversion (0.005), agreeableness (0.021), and neuroticism (−0.042). The personality traits for which we detect effects of additional day care attendance are only those that are more correlated with IQ.[50]

The ITT results translate into the estimates of the effect of an additional month in day care presented in table 9. For the more affluent, this treatment decreases openness and agreeableness by 1.4% and 1.2%, respectively, and increases neuroticism by 0.9%, with no effect for the less affluent and for the other traits. A comparison with the IQ effect for the more affluent (−0.9%) suggests that cognitive and noncognitive skills respond similarly at age 8–14 to a substitution of informal care with formal care in households characterized by different socioeconomic status and therefore different quality of informal care. The similarities in the effects for the cognitive and noncognitive domains is not surprising given that these skills are similarly sensitive to early influences in life, as discussed in section VII.

Table 10 replicates for the big five personality traits the IV analysis by cases L and N reported in table 7 for IQ, with qualitatively similar results. Specifically, we observe no statistically significant difference between the L and N cases.

VII. Suggestions from the Psychological Literature

Psychologists have produced persuasive empirical evidence that during the first 3 years of life, one-to-one interactions with adults (more than interactions with peers) are a crucial input for both the cognitive and the noncognitive development of a child. For instance, in an empirical field study of 42 American families, Hart and Risley (1995) have recorded one

[50] With specific reference to conscientiousness, this pattern is in line with a remark in Elango et al. (2016, 254): this personality trait is “a non-cognitive skill that is of interest due to its low correlation with cognition and high correlation with important later-life outcomes.”

TABLE 10
EFFECTS OF DAY CARE 0–2 ATTENDANCE ON PERSONALITY, BY CASES L AND N

	All Thresholds		≤Median		>Median	
	β_L	$\beta_N - \beta_L$	β_L	$\beta_N - \beta_L$	β_L	$\beta_N - \beta_L$
Openness	−.004	.000	.003	−.010	−.019*	.009
	(.006)	(.007)	(.009)	(.011)	(.009)	(.010)
Conscientiousness	−.001	.002	.011	−.004	−.007	.012
	(.006)	(.006)	(.008)	(.012)	(.010)	(.010)
Extraversion	−.007	−.000	−.014	.010	−.003	−.007
	(.007)	(.007)	(.010)	(.012)	(.009)	(.011)
Agreeableness	−.006	.003	.004	.007	−.015	.006
	(.006)	(.008)	(.008)	(.008)	(.009)	(.013)
Neuroticism	.002	−.001	−.009	.010	.016$^+$	−.012
	(.006)	(.007)	(.010)	(.009)	(.009)	(.009)
Observations	447	447	225	225	222	222

NOTE.—The table reports parametric IV estimates of the effect of 1 month of day care 0–2 on the log scores in the BFQ-C, at all levels of the preferred FAI threshold and separately for preferred FAI thresholds below or above the median preferred FAI threshold. Coefficients are from regressions of each outcome on months of attendance, months of attendance interacted with $\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M)$ and the full set of controls A_i and X_i using the dummy P_i for qualification in the preferred program, and the same dummy interacted with $\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M) = 1 - \Omega$ as the instruments (see n. 48). The running variable is the FAI, and the polynomial in the running variable is of second order. Sample: 447 (446 for agreeableness) interviewed children with two working parents, born between 1999 and 2005, with nonmissing outcome or covariates and whose parents first applied for admission to day care between 2001 and 2005. Robust standard errors are shown in parentheses, clustered at the facility level.
$^+$ Significant at 10%.
* Significant at 5%.

full hour of words spoken at home every month for 2.5 years by parents with their children at age 0–2. They conclude that "the size of the children's recorded vocabularies and their IQ scores were strongly associated with the size of their parents' recorded vocabulary and their parents' scores on a vocabulary pre-test" (Hart and Risley 1995, 176). Along the same lines, Rowe and Goldin-Meadow (2009) and Cartmill et al. (2013) show that the quality of parental inputs in the first 3 years of life (e.g., in terms of parental gesture and talking) improves children's vocabulary before entering school. Similarly, Gunderson et al. (2013) finds that parental praise directed to 1- to 3-year-old children predicts their motivation 5 years later.[51]

A fascinating theory explaining why early one-to-one interactions with adults are so important has been proposed by Csibra and Gergely (2009, 2011). According to these authors, the communication between a trusted

[51] Related to these results, some psychologists have estimated negative effects of increasing parental (in particular maternal) working time on cognitive and noncognitive outcomes of children. See, e.g., Brooks-Gunn, Han, and Waldfogel (2002); Adi-Japha and Klein (2009); Lombardi and Coley (2014); and the meta-analysis in Li et al. (2013). Unlike the economic literature, however, most of these studies are observational and do not exploit quasi-experimental identification strategies.

adult and a child allows the latter to understand more rapidly whether an experience has a general value or only a specific one. Lacking such communication, the child must repeat an experience many times to assess its general or particular validity (very much like statistical inference requiring a large sample). An adult, instead, can quickly inform the child about the nature of what he or she is experimenting. If the adult can be trusted, then the child can save time and move on to other experiences, thus gaining a developmental advantage.

The focus on one-to-one interactions in our context is relevant because, as noted by Clarke-Stewart, Gruber, and Fitzgerald (1994), infants and toddlers generally experience less one-to-one attention in day care than at home because at home they are typically taken care of by a parent, a grandparent, or a nanny. Under these care modes, a child receives attention in a 1:1 ratio, possibly somewhat less if, for example, siblings are present. This is precisely the case for the children in our sample. When we asked their parents which options were available at the time of the first application as an alternative to day care during the workday, 50.5% checked "the mother," 11% "the father," 44.8% "the grandparents," 4.5% "other family members," 18.9% "a babysitter or a nanny," and only 12.1% checked "some other day care center" (multiple answers were possible).[52] The adult-to-child ratio in day care 0–2 depends instead on the specific institutional setting. At the BDS, during the period under investigation, this ratio was 1:4 at age 0 and 1:6 at age 1–2. This may be part of the reason why both Felfe and Lalive (2018) and Drange and Havnes (2019) find positive effects of day care 0–2 in Germany and Norway. In their institutional setting, the adult-to-child ratio is about 1:3.

A related hypothesis emphasized by Belsky and Steinberg (1978) and Belsky (1988, 2001) is that the negative effects of day care are driven by decreased interactions with the mother, leading to a drop in maternal involvement, children's attachment to their mothers, and consequent child insecurity. This maternal channel is at center stage in Bernal (2008). We cannot tell whether the negative effects of day care that we uncover are driven by a substitution away from maternal care or from family-based

[52] In Bologna, there are very few private day care facilities outside the public system. The reason is that Bologna is one of the Italian cities with the largest and most highly reputed public day care systems, which leaves little room for independent private providers, relative to other cities. The BDS includes nine charter facilities that are privately managed but strictly regulated by the BDS. According to the reputational indicator of quality described in sec. IV.A, these charter programs are perceived by parents as worse than the noncharter ones. On average, for given distance, charter programs are ranked 1.6 positions lower than the noncharter ones if they are included in the application set of a parent. Moreover, the probability that a charter program is not ranked by parents in their application set is higher than for noncharter ones (0.95 vs. 0.91). The odds that a charter program is ranked first by a parent is 0.007, while the odds that a program is charter is 0.046. Therefore, it is unlikely that the worse quality of these charter programs is responsible for the negative effects that we estimate—which derive from the offer of the most preferred program to a parent. If anything, their presence should reduce the absolute size of our estimates.

care more generally. However, the aforementioned statistics indicate that in our sample the counterfactual care mode for the fraction of time children would not have spent in day care is not only mothers.

Psychologists have also supported with persuasive empirical evidence the idea that girls are more "mature" than boys, in the sense of being more capable of absorbing cognitive stimuli at an early age. For example, Fenson et al. (1994) study 1,800 toddlers (16–30 months of age), finding that girls perform better in lexical, gestural, and grammatical development. Galsworthy et al. (2000) examine about 3,000 2-year-old twin pairs and show that girls score higher on verbal and nonverbal tests. A longitudinal study by Bornstein, Hahn, and Hayne (2004, 206) involving 329 children observed between ages 2 and 5 reaches similar conclusions for an age range partially overlapping with ours: they show that "girls consistently outperformed boys in multiple specific and general measures of language."

If girls at age 0–2 are relatively more capable of making good use of stimuli that improve their skills, then their development is hurt (more than for boys) by an extended exposure to day care because it implies fewer one-to-one interactions with adults, which are more valuable for them than for boys as inputs in the technology of skill formation. In the appendix to section III, we extend our theoretical framework to model how parental decisions concerning childcare are compatible with these gender differences in the effects of day care 0–2. A replication of our analysis by gender for both IQ and the big five personality traits is also reported in the appendix to section VII and can be summarized as follows: for girls, the ITT effects of barely qualifying for the most preferred program consist of a reduction in IQ by 3.9% (*p*-value: .017) and extraversion by 7.2% (*p*-value: .036), with a decrease in openness of 10.5% (*p*-value: .053) that emerges for more affluent girls. For boys, the corresponding coefficients are smaller, and we cannot reject the hypothesis that they are equal to zero.[53] A similar gender difference also emerges from the IV estimates, which indicate that for girls one additional month in day care 0–2 reduces IQ by 0.7% (*p*-value: .016) and, when restricting to more affluent girls, reduces openness by 1.8% (*p*-value: .067) and extraversion by 2% (*p*-value: .097), with smaller and insignificant effects for boys.[54] This gender heterogeneity in the cognitive and noncognitive losses induced by day care attendance supports the relevance of one-to-one interactions with adults as an explanation of our results.

[53] This gender gap is not due to differences in the first stage. Similarly, we can easily dismiss the possibility that this gender heterogeneity in the effects of day care 0–2 reflects differences in pretreatment characteristics of boys and girls in our sample. The appendix reproduces the figure and tables of the main text related to the validity of our identification strategy separately for the two gender groups.

[54] Interestingly, in a longitudinal study of 113 first-born preschool children (58 girls and 55 boys), Bornstein et al. (2006, 145) find, in line with our results, that "girls who had greater amount of non-maternal care from birth to 1 year scored lower on the Spoken Language Quotient at preschool."

VIII. Alternative Mechanisms

We have also explored some alternative interpretations of our results. A first possibility is that day care 0–2 may negatively affect IQ because it increases children's exposure to infections (Baker, Gruber, and Milligan 2008), which have been shown to harm human capital accumulation and cognitive development (see Eppig, Fincher, and Thornill 2010; John, Black, and Nelson 2017; and the review in Bleakley 2010). However, since boys are more vulnerable than girls to infection exposure at a young age (Muenchhoff and Goulder 2014), this explanation is at odds with the gender difference in the effects of day care 0–2 that we uncover.

Second, in line with the early results reported in Belsky and Steinberg (1978) and additional findings in Baker, Gruber, and Milligan (2008), it is also possible that day care 0–2 induces a disengagement of parents from the education of children, which may have a negative effect on their development. We do not have direct evidence to dismiss this interpretation, but, again in light of the gender heterogeneity of our results, it is not clear why parental disengagement should be more pronounced for girls than for boys.

A third alternative mechanism that might specifically explain the gender difference that we estimate in the effects of day care 0–2 refers to the possibility that the loss suffered by girls depends on the sex ratio within each program. Psychologists have observed that in early education, "teachers spend more time socializing boys into classroom life, and the result is that girls get less teacher attention. Boys receive what they need. . . . Girls' needs are more subtle and tend to be overlooked" (Koch 2003, 265). However, we do not find any evidence that sex ratios affect the size of the effects for girls and boys, perhaps because the variation in these ratios is quite small for the children in our sample.

Finally, the data do not support the hypothesis that gender differences in breastfeeding explain the gender gap in the effects of day care. The duration of breastfeeding has been shown to be positively associated with cognitive outcomes (Anderson, Johnstone, and Remley 1999; Borra, Iacovou, and Sevilla 2012; Fitzsimons and Vera-Hernandez 2013), and early day care attendance may shorten it. However, we find no effect (and specifically no differential effect by gender) of days in day care on the duration of breastfeeding.

IX. Conclusions

This paper contributes to a growing literature studying the effects of time spent in day care 0–2 on child ability. We studied the offspring of dual-earner households with cohabiting parents in Bologna, one of the most educated and richest Italian cities with a highly reputed public day care system. For the children in this population, our results indicate a quantitatively and statistically significant loss in IQ and some noncognitive traits at age 8–14. This loss is even more pronounced when we focus on children with relatively more affluent parents within this population. These are typically the relevant marginal subjects to be considered in an evaluation of day care expansions as a response to the growing incidence of dual-earner households in advanced countries.

We interpret these findings in a theoretical model showing that, when offered their most preferred day care program (as opposed to a less preferred one or no offer), relatively affluent parents take advantage of this opportunity to increase labor supply and child's day care attendance. Because of the higher earning potential of more affluent parents, this increase in attendance generates an increase of family resources that is large enough to become attractive even if it comes at the cost of child ability.

These results seem relevant not only because of their novelty with respect to the literature but also because they implicitly support the hypothesis, suggested by psychologists, that the sign and size of the effects of day care 0–2 are mostly driven by three factors: first, whether this early life experience deprives children of one-to-one interactions with adults at home; second, the quality of these interactions, which is likely to be higher in more affluent households; and third, whether children can make good use of these interactions. The latter claim is supported by the finding that day care 0–2 has a more negative effect on the ability of girls, in combination with the psychological evidence suggesting that girls are developed enough at this young age to exploit higher-quality interactions with adults that are less valuable for boys.

Our identification strategy exploits affluence thresholds that discriminate between similar parents whose children attend day care 0–2 for longer versus shorter periods because they are barely admitted to their preferred program instead of being simply excluded from it. This strategy makes our results valuable not only for parents but also for policy makers interested in expanding vacancies in the day care systems that they manage. The estimates we presented speak precisely about the effect of such a policy, which would allow more affluent children to attend for a longer time in programs that their parents prefer more, with negative effects on their skills that may not be socially optimal even if the utility of parents increase.

References

Adi-Japha, Esther, and Pnina S. Klein. 2009. "Relations between Parenting Quality and Cognitive Performance of Children Experiencing Varying Amounts of Childcare." *Child Development* 80 (3): 893–906.

Almond, Douglas, and Janet Currie. 2011. "Human Capital Development before Age Five." In *Handbook of Labor Economics*, vol. 4, pt. B, edited by Stephen Machin, Eric A. Hanushek, and Ludger Woessmann, 1315–486. Amsterdam: North-Holland.

Anderson, James W., Bryan M. Johnstone, and Daniel T. Remley. 1999. "Breast-Feeding and Cognitive Development: A Meta-analysis." *American J. Clinical Nutrition* 70 (4): 525–35.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *J. American Statis. Assoc.* 103 (484): 1481–95.

Angrist, Joshua, and Guido Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *J. American Statis. Assoc.* 90 (430): 431–42.

Baker, Michael, Jonathan Gruber, and Kevin Milligan. 2008. "Universal Child Care, Maternal Labor Supply, and Family Well-Being." *J.P.E.* 116 (4): 709–45.

———. 2015. "Non-cognitive Deficits and Young Adult Outcomes: The Long-Run Impacts of a Universal Child Care Program." Working Paper no. 21571, NBER, Cambridge, MA.

Barrett, Garry F., and Stephen G. Donald. 2003. "Consistent Tests for Stochastic Dominance." *Econometrica* 71 (1): 71–104.

Becker, Gary S. 1965. "A Theory of the Allocation of Time." *Econ. J.* 75 (299): 493–517.

Belsky, Jay. 1988. "The 'Effects' of Infant Day Care Reconsidered." *Early Childhood Res. Q.* 3 (3): 235–72.

———. 2001. "Emanuel Miller Lecture: Developmental Risks (Still) Associated with Early Child Care." *J. Child Psychology and Psychiatry* 42 (7): 845–59.

Belsky, Jay, and Laurence D. Steinberg. 1978. "The Effects of Day Care: A Critical Review." *Child Development* 49 (4): 929–49.

Bernal, Raquel. 2008. "The Effect of Maternal Employment and Child Care on Children's Cognitive Development." *Internat. Econ. Rev.* 49 (4): 1173–209.

Bernal, Raquel, and Michael P. Keane. 2011. "Child Care Choices and Children's Cognitive Achievement: The Case of Single Mothers." *J. Labor Econ.* 29 (3): 459–512.

Bertanha, Marinho. 2017. "Regression Discontinuity Design with Many Thresholds." Working paper, Dept. Econ., Univ. Notre Dame. https://doi.org/10.2139/ssrn.2712957.

Bleakley, Hoyt. 2010. "Health, Human Capital, and Development." *Ann. Rev. Econ.* 2:283–310.

Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas ter Weel. 2008. "The Economics and Psychology of Personality Traits." *J. Human Resources* 43 (4): 972–1059.

Bornstein, Marc, Chun-Shin Hahn, Nancy Gist, and Maurice Haynes. 2006. "Long-Term Cumulative Effects of Childcare on Children's Mental Development and Socioemotional Adjustment in a Non-risk Sample: The Moderating Effects of Gender." *Early Child Development and Care* 176 (2): 129–56.

Bornstein, Marc H., Chun-Shin Hahn, and O. Maurice Haynes. 2004. "Specific and General Language Performance across Early Childhood: Stability and Gender Considerations." *First Language* 24 (3): 267–304.

Borra, Cristina, Maria Iacovou, and Almudena Sevilla. 2012. "The Effect of Breastfeeding on Children's Cognitive and Noncognitive Development." *Labour Econ.* 19 (4): 496–515.

Brooks-Gunn, Jeanne, Wen-Jui Han, and Jane Waldfogel. 2002. "Maternal Employment and Child Cognitive Outcomes in the First Three Years of Life: The NICHD Study of Early Child Care." *Child Development* 73 (4): 1052–72.

Calonico, Sebastian, Matias D. Cattaneo, and Max H. Farrell. 2018. "On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference." *J. American Statis. Assoc*. 113 (522): 767–79.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocío Titiunik. 2019. "Regression Discontinuity Designs Using Covariates." *Rev. Econ. and Statis*. 101 (3): 442–51.

Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82 (6): 2295–326.

Campbell, Frances A., and Craig T. Ramey. 1994. "Effects of Early Intervention on Intellectual and Academic Achievement: A Follow-Up Study of Children from Low-Income Families." *Child Development* 65 (2): 684–98.

Canay, Ivan A., and Vishal Kamat. 2018. "Approximate Permutation Tests and Induced Order Statistics in the Regression Discontinuity Design." *Rev. Econ. Studies* 85 (3): 1577–608.

Card, David, David S. Lee, Zhuan Pei, and Andrea Weber. 2015. "Inference on Causal Effects in a Generalized Regression Kink Design." *Econometrica* 83 (6): 2453–83.

Carneiro, Pedro, Flavio Cunha, and James Heckman. 2003. "Interpreting the Evidence of Family Influence on Child Development." Paper presented at the Economics of Early Childhood Development: Lessons for Economic Policy conference, Minneapolis, October 17.

Carneiro, Pedro, and Rita Ginja. 2014. "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start." *American Econ. J.: Econ. Policy* 6 (4): 135–73.

Carneiro, Pedro, Katrine V. Løken, and Kjell G. Salvanes. 2015. "A Flying Start? Maternity Leave Benefits and Long-Run Outcomes of Children." *J.P.E*. 123 (2): 365–412.

Cartmill, Erica, Benjamin Armstrong, Lila Gleitman, Susan Goldin-Meadow, Tamara Medina, and John Trueswell. 2013. "Quality of Early Parent Input Predicts Child Vocabulary 3 Years Later." *Proc. Nat. Acad. Sci. USA* 110 (28): 11278–83.

Cascio, Elizabeth U., and Diane Whitmore Schanzenbach. 2013. "The Impacts of Expanding Access to High-Quality Preschool Education." *Brookings Papers Econ. Activity* 47 (2): 127–92.

Cattaneo, Matias D., Brigham R. Frandsen, and Rocío Titiunik. 2015. "Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate." *J. Causal Inference* 3 (1): 1–24.

Cattaneo, Matias D., Luke Keele, Rocío Titiunik, and Gonzalo Vazquez-Bare. 2016. "Interpreting Regression Discontinuity Designs with Multiple Cutoffs." *J. Polit*. 78 (4): 1229–48.

Clarke-Stewart, K. Alison, Christian Gruber, and Linda May Fitzgerald. 1994. *Children at Home and in Day Care*. Hillsdale, NJ: Erlbaum.

Comune di Bologna. 2010. "Progetto Pedagogico del Nido d'Infanzia del Comune di Bologna." Bologna: Comune di Bologna Settore Istruzione.

Csibra, Gergely, and Gyorgy Gergely. 2009. "Natural Pedagogy." *Philosophical Transactions Royal Soc. B* 13 (4): 148–53.

———. 2011. "Natural Pedagogy as Evolutionary Adaptation." *Trends Cognitive Sci.* 366:1149–57.

Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *A.E.R.* 97 (2): 31–47.

Currie, Janet. 2001. "Early Childhood Education Programs." *J. Econ. Perspectives* 15 (2): 213–38.

de la Cuesta, Brandon, and Kosuke Imai. 2016. "Misunderstandings about the Regression Discontinuity Design in the Study of Close Elections." *Ann. Rev. Polit. Sci*. 19:375–96.

Drange, Nina, and Tarjei Havnes. 2019. "Child Care before Age Two and the Development of Language and Numeracy: Evidence from a Lottery." *J. Labor Econ*. 37 (2): 581–620.

Duncan, Greg J., and Katherine Magnuson. 2013. "Investing in Preschool Programs." *J. Econ. Perspectives* 27 (2): 109–32.

Duncan, Greg J., and Aaron J. Sojourner. 2013. "Can Intensive Early Childhood Intervention Programs Eliminate Income-Based Cognitive and Achievement Gaps?" *J. Human Resources* 48 (4): 945–68.

Elango, Sneha, Jorge Luis García, James J. Heckman, and Andrés Hojman. 2016. "Early Childhood Education." In *Economics of Means-Tested Transfer Programs in the United States*, vol. 2, edited by Robert A. Moffitt, 235–97. Chicago: Univ. Chicago Press.

Eppig, Christopher, Corey L. Fincher, and Randy Thornill. 2010. "Parasite Prevalence and the Worldwide Distribution of Cognitive Ability." *Proc. Royal Soc. B* 277:3801–8.

Felfe, Christina, and Rafael Lalive. 2018. "Does Early Child Care Help or Hinder Children's Development?" *J. Public Econ*. 159:33–53.

Felfe, Christina, Natalia Nollenberger, and Núria Rodríguez-Planas. 2015. "Can't Buy Mommy's Love? Universal Childcare and Children's Long-Term Cognitive Development." *J. Population Econ*. 28 (2): 393–422.

Fenson, Larry, Philip S. Dale, J. Steven Reznick, Elizabeth Bates, Donna J. Thal, Stephen J. Pethick, Michael Tomasello, Carolyn B. Mervis, and Joan Stiles. 1994. "Variability in Early Communicative Development." *Monographs Soc. Res. Child Development* 59 (5): 1–173.

Fiorini, Mario, and Katrien Stevens. 2017. "Assessing the Monotonicity Assumption in IV and Fuzzy RD Designs." Working Paper no. 2014-13, School Econ., Univ. Sydney.

Fitzsimons, Emla, and Marcos Vera-Hernandez. 2013. "Food for Thought? Breastfeeding and Child Development." IFS Working Paper no. W13/31, Inst. Fiscal Studies, London.

Florens, J. P., J. J. Heckman, C. Meghir, and E. Vytlacil. 2008. "Identification of Treatment Effects Using Control Functions in Models with Continuous, Endogenous Treatment and Heterogeneous Effects." *Econometrica* 76 (5): 1191–206.

Gale, David, and Lloyd Shapley. 1962. "College Admissions and the Stability of Marriage." *American Math. Monthly* 69:9–15.

Galsworthy, Michael J., Ginette Dionne, Philip S. Dale, and Robert Plomin. 2000. "Sex Differences in Early Verbal and Non-verbal Cognitive Development." *Developmental Sci*. 3 (2): 206–15.

Garber, Howard. 1988. "The Milwaukee Project: Preventing Mental Retardation in Children at Risk." Technical report, American Assoc. Mental Retardation, Washington, DC.

Gormley, William T., and Ted Gayer. 2005. "Promoting School Readiness in Oklahoma: An Evaluation of Tulsa's Pre-K Program." *J. Human Resources* 40 (3): 533–58.

Gottfredson, Linda S. 1997. "Mainstream Science on Intelligence: An Editorial with 52 Signatories, History, and Bibliography." *Intelligence* 24 (1): 13–23.

Gunderson, Elizabeth A., Sarah J. Gripshover, Carissa Romero, Carol S. Dweck, Susan Goldin-Meadow, and Susan C. Levine. 2013. “Parent Praise to 1- to 3-Year-Olds Predicts Children’s Motivational Frameworks 5 Years Later.” *Child Development* 84 (5): 1526–41.

Hart, Betty, and Todd Risley. 1995. *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore: Brookes.

Havnes, Tarjei, and Magne Mogstad. 2015. “Is Universal Child Care Leveling the Playing Field?” *J. Public Econ*. 127:100–114.

Heckman, James J., and Stefano Mosso. 2014. “The Economics of Human Development and Social Mobility.” *Ann. Rev. Econ*. 6:689–733.

Herbst, Chris M. 2013. “The Impact of Non-parental Child Care on Child Development: Evidence from the Summer Participation ‘Dip.’ ” *J. Public Econ*. 105:86–105.

Hewett, Valarie Mercilliott. 2001. “Examining the Reggio Emilia Approach to Early Childhood Education.” *Early Childhood Educ. J*. 29 (2): 95–100.

Imbens, Guido W., and Thomas Lemieux. 2008. “Regression Discontinuity Designs: A Guide to Practice.” *J. Econometrics* 142 (2): 615–35.

John, Chandy C., Maureen M. Black, and Charles A. Nelson. 2017. “Neurodevelopment: The Impact of Nutrition and Inflammation during Early to Middle Childhood in Low Resource Settings.” *Pediatrics* 139 (suppl.): S59–S71.

Koch, Janice. 2003. “Gender Issues in the Classroom.” In *Educational Psychology*, vol. 7 of *Handbook of Psychology*, edited by William M. Reynolds and Gloria E. Miller, 259–84. Hoboken, NJ: Wiley.

Kottelenberg, Michael J., and Steven F. Lehrer. 2014a. “Do the Perils of Universal Childcare Depend on the Child’s Age?” *CESifo Econ. Studies* 60 (2): 338–65.

———. 2014b. “The Gender Effects of Universal Child Care in Canada: Much Ado about Boys?” Working paper, Queen’s Univ., Kingston, Ontario.

———. 2017. “Targeted or Universal Coverage? Assessing Heterogeneity in the Effects of Universal Childcare.” *J. Labor Econ*. 35 (3): 609–53.

Lee, David S., and Thomas Lemieux. 2010. “Regression Discontinuity Designs in Economics.” *J. Econ. Literature* 48 (2): 281–355.

Li, Jianghong, Sarah E. Johnson, Wen-Jui Han, Sonia Andrews, Garth Kendall, Lyndall Strazdins, and Alfred Dockery. 2013. “Parents’ Nonstandard Work Schedules and Child Well-Being: A Critical Review of the Literature.” *J. Primary Prevention* 35 (1): 53–73.

Lombardi, Caitlin McPherran, and Rebekah Levine Coley. 2014. “Early Maternal Employment and Children’s School Readiness in Contemporary Families.” *Developmental Psychology* 50 (8): 2071–84.

McCrary, Justin. 2008. “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test.” *J. Econometrics* 142 (2): 698–714.

Muenchhoff, Maximilian, and Patrick J. R. Goulder. 2014. “Sex Differences in Pediatric Infectious Diseases.” *J. Infectious Diseases* 209 (3): S120–S126.

Noboa-Hidalgo, Grace E., and Sergio S. Urzùa. 2012. “The Effects of Participation in Public Child Care Centers: Evidence from Chile.” *J. Human Capital* 6 (1): 1–34.

Puma, Mike, Stephen Bell, Ronna Cook, Camilla Heid, Pam Broene, Frank Jenkins, Andrew Mashburn, and Jason Downer. 2012. “Third Grade Follow-Up to the Head Start Impact Study Final Report.” OPRE Report no. 2012-45, Office Planning Res. and Evaluation, Admin. Children and Families, US Dept. Health and Human Services, Washington, DC.

Roth, Alvin. 2008. “Deferred Acceptance Algorithms: History, Theory, Practice, and Open Questions.” *Internat. J. Game Theory* 36 (3): 537–69.

Rowe, Meredith L., and Susan Goldin-Meadow. 2009. "Differences in Early Gesture Explain SES Disparities in Child Vocabulary Size at School Entry." *Science* 323 (5916): 951–53.

Weiland, Christina, and Hirokazu Yoshikawa. 2013. "Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills." *Child Development* 84 (6): 2112–30.

Zigler, Edward, and Earl C. Butterfield. 1968. "Motivational Aspects of Changes in IQ Test Performance of Culturally Deprived Nursery School Children." *Child Development* 39 (1): 1–14.